

UNIVERSIDADE FEDERAL DE SERGIPE

PÓS-GRADUAÇÃO EM CIÊNCIA DA COMPUTAÇÃO

Alinhamento do Modelo de Forma Ativa com Máquinas de Vetores de Suporte Aplicado na Detecção de Veículos

Maria Gêssica dos Santos Aragão

São Cristóvão

2016

Maria Gécica dos Santos Aragão

Alinhamento do Modelo de Forma Ativa com Máquinas de Vetores de Suporte Aplicado na Detecção de Veículos

Dissertação apresentada ao Programa de
Pós-graduação em Ciência da Computação
da Universidade Federal de Sergipe como
requisito parcial para obtenção do grau de
Mestre em Ciência da Computação.

Orientador: Prof. Dr. Leonardo Nogueira
Matos.

São Cristóvão
2016

**FICHA CATALOGRÁFICA ELABORADA PELA BIBLIOTECA CENTRAL
UNIVERSIDADE FEDERAL DE SERGIPE**

A659a Aragão, Maria Géssica dos Santos
 Alinhamento do modelo de forma ativa com máquinas de
 vetores de suporte aplicado na detecção de veículos / Maria Géssica
 dos Santos Aragão; orientador Leonardo Nogueira Matos. – São
 Cristóvão, 2016.
 55 f.: il.

 Dissertação (Mestrado em Ciência da Computação) -
 Universidade Federal de Sergipe, 2016.

 1. Processamento de imagens – Técnicas digitais. 2. Campos
 vetoriais. 3. Sistema de computação. I. Matos, Leonardo Nogueira.
 II. Título

CDU: 004.932



UNIVERSIDADE FEDERAL DE SERGIPE
PRÓ-REITORIA DE PÓS-GRADUAÇÃO E PESQUISA
NÚCLEO DE PÓS-GRADUAÇÃO EM CIÊNCIA DA COMPUTAÇÃO

Relatório de defesa pública do(a) Senhor(a) **MARIA GÉSSICA DOS SANTOS ARAGÃO** no Programa de Ciência da Computação (PROCC) da UFS.

Aos 13 dias do mês de maio de 2016, realizou-se a **Defesa de Mestrado** do trabalho intitulado "**Alinhamento do Modelo de Forma Ativa com Maquinas de Vetores de Suporte Aplicado na Detecção de Veículos**" sob orientação do Prof. Dr. **Leonardo Nogueira Matos**.

Depois de declarada aberta a sessão, o Presidente da Banca passou inicialmente a palavra ao candidato para exposição e a seguir aos examinadores para as devidas arguições que se desenvolveram nos termos regimentais. Em seguida, a comissão julgadora proclamou o resultado:

Nome Banca Examinadora	Instituição	Assinatura
Leonardo Nogueira Matos	UFS	
Eduardo Oliveira Freire	UFS	
Adilson Gonzaga	USP	

Dessa maneira o Resultado Final é:



APROVADO ou

Reprovado

Parecer da Banca Examinadora *

- Obs: Se o candidato for reprovado, o preenchimento do parecer é obrigatório.

São Cristóvão/SE

Assinatura do Orientador:

Assinatura do Aluno:

Resumo

Muitas aplicações de processamento de imagens digitais utilizam técnicas de detecção de objetos. Detetar um objeto normalmente está relacionado a localizar a área em torno do mesmo, enquanto a detecção da forma está relacionada a localizar, precisamente, o conjunto de pontos que constituem sua forma. Quando o problema envolve a detecção de formas que apresentam variações previsíveis, os modelos deformáveis mostram-se uma solução eficaz. A abordagem desenvolvida neste trabalho se refere à detecção da forma de veículos em posição frontal através de métodos que se dividem em dois níveis, o primeiro nível é composto por uma cascata de máquinas de vetores de suporte e o segundo é um modelo deformável. O uso de modelos deformáveis favorece a detecção de formas de veículos mesmo quando sua imagem está ocluída por objetos, tais como árvores.

Palavras-chave: Segmentação da forma de veículos, Máquinas de Vetores de Suporte, Modelos Deformáveis, Modelos de Forma Ativa.

Abstract

Many applications of digital image processing uses object detection techniques. Detecting an object is usually related to locate the area around it, while shape detection is related to find, precisely, the set of points that constitutes its shape. When the problem involves detecting shapes that have predictable changes, deformable models show to be an effective solution. The approach developed in this work refers to the vehicle shape detection in frontal position by methods which are divided into two levels, the first level is composed by a cascade of support vector machines and the second one is a deformable model. The use of deformable models favors the detection of vehicle shape same when its image is occluded by objects such as trees

Keywords: Vehicle shape segmentation, Support Vector Machines, Deformable Models, Active Shape Models.

Lista de Figuras

1.1	Visão geral da abordagem	10
2.1	Pontos de controle	13
2.2	Alinhamento de amostras através do algoritmo Análise Generalizada de Procrustes; (a) amostras desalinhadas, (b) amostras alinhadas	16
2.3	Resultado da Análise de Componentes Principais	17
2.4	Busca por perfis de níveis de cinza	19
3.1	Deslocamento associado ao ponto de controle	22
3.2	Modelagem com AAM	23
3.3	Visão geral do sistema, (a) imagem de entrada, (b) alinhamento do modelo 3D para imagem 2D, (c) Correção dos veículos para o mesmo ponto de vista e extração de partes dos veículos (por exemplo, motor, rodas, e farol) e (d) 5 melhores resultados de pesquisa através da fusão das três partes	24
3.4	Características dos pontos de controle em diferentes visões do veículo, (a) semi frontal, (b) perfil, (c) semi fundo, (d) frontal e (e) fundo	25
3.5	Ilustração do processo de detecção de veículos	26
4.1	Exemplo de amostras positivas e negativas para treinamento do Modelo de Forma Ativa	28
4.2	Amostra de treinamento	28
4.3	Resultados da classificação da Máquina de Vetores de Suporte; (a) – (f) Mapa de saliência obtido para cada ponto	29
4.4	Estimativa inicial	30
4.5	Características HOG de cada ponto	31
4.6	Diagrama da visão geral dos experimentos	32
5.1	Diagrama do Experimento I	34
5.2	Diagrama do Experimento II	35
5.3	Diagrama do Experimento III	36
5.4	Diagrama do Experimento IV	37

5.5	Amostras da base MIT Street Scenes	38
5.6	Histograma do mapa produzido pela SVM de regressão	39
5.7	Resultados ASM em todos os experimentos	43
5.8	Resultados ASM em veículos oclusos	44
5.9	Erros dos experimentos I ao IV	45
5.10	Erros da abordagem de Li, Gu e Kanade [2011] com vista frontal	45

Lista de Tabelas

5.1	Experimento I - Medidas de desempenho das SVM no primeiro nível . . .	39
5.2	Experimento II - Medidas de desempenho das SVM no segundo nível com LBP	40
5.3	Experimento III - Medidas de desempenho das SVM no segundo nível com HOG	40
5.4	Experimento IV - Medidas de desempenho das RF no primeiro nível . . .	40
5.5	Medidas de desempenho do ASM	41

Lista de Símbolos

AAM	<i>Active Appearance Model</i>
ASM	<i>Active Shape Model</i>
GPU	<i>Graphics Processing Unit</i>
HOG	<i>Histogram of Oriented Gradients</i>
ITS	<i>Intelligent Transport Systems</i>
LBP	<i>Local Binary Pattern</i>
MIT	<i>Massachusetts Institute of Technology</i>
PDM	<i>Points Distribution Model</i>
PR	<i>Point Registration</i>
RF	<i>Random Forest</i>
RMSE	<i>Root Mean Squared Error</i>
SVM	<i>Support Vector Machine</i>
SVR	<i>Support Vector Regression</i>
UCI	<i>University of California Irvine</i>

Sumário

1	Introdução	7
1.1	Motivação	8
1.2	Objetivos	9
1.3	Metodologia	9
1.4	Estrutura da Dissertação	10
2	Modelos Pontuais	12
2.1	Modelo Pontual de Distribuição (PDM)	12
2.1.1	Etiquetagem	13
2.1.2	Alinhamento (Análise Generalizada de Procrustes)	13
2.1.3	Variações Admissíveis	15
2.1.4	Perfis de Intensidade	17
2.2	Busca da Forma na Imagem (Modelo de Forma Ativa)	18
2.2.1	Estimativa Inicial	18
2.2.2	Cálculo do Movimento Sugerido para Cada Ponto do Modelo	19
2.2.3	Computação das Alterações dos Parâmetros de Pose e Forma	20
2.2.4	Atualização dos Parâmetros	20
3	Modelos Pontuais e Detecção de Veículos	22
4	Desenvolvimento da Abordagem	27
4.1	Base de Imagens	27
4.2	Modelo de Forma Ativa	29
4.2.1	Estimativa Inicial	29
4.2.2	Cálculo dos Deslocamentos Através do Classificador SVM	29
5	Experimentos e Resultados	33
5.1	Avaliação dos Classificadores SVM	38
5.2	Avaliação do Modelo de Forma Ativa	41

6	Conclusões	46
6.1	Trabalhos Futuros	47

Capítulo 1

Introdução

A detecção de veículos é um problema bastante desafiador na Visão Computacional. Devido à grande variação intra-classe, diferentes pontos de vista, oclusões, iluminação e fundo complexo, a detecção de veículos transforma-se em um problema ainda mais difícil. Para objetos com menor variação intra-classe, como faces frontais, há métodos bem sucedidos, como o classificador em cascata de Viola e Jones [2001]. No entanto, para uma classe de objetos com grande variação intra-classe é necessário empregar métodos mais complexos, como discutido em Kuo e Nevatia [2009].

Durante as duas décadas passadas, modelos de formas deformáveis, do inglês *deformable models* [Cootes e Taylor, 2004], vêm sendo objeto de estudo pela comunidade de Visão Computacional. O primeiro destes modelos deve-se a Kass, Witkin e Terzopoulos [1988], denominado Modelo de Contornos Ativos, que consiste em um conjunto de pontos que se movem para se ajustar à forma de objetos de interesse em imagens.

Os Modelos Pontuais de Distribuição, do inglês *Points Distribution Model* — PDM surgiram como um método mais robusto, pois utilizam técnicas estatísticas para estabelecer o intervalo de variação para uma forma encontrada na imagem. Dentre as variações do modelo pontual, há o Modelo de Forma Ativa, do inglês *Active Shape Model* — ASM, proposto por Cootes et al. [1995].

Vários autores têm utilizado modelos pontuais na detecção de veículo, dentre esses modelos o Modelo de Forma Ativa. A utilização dos modelos deformáveis nos Sistemas de Transportes Inteligentes se aplica em diferentes situações como detecção de automóveis em imagens do centro da cidade [Zia, Stark e Schindler, 2013], detecção de veículos oclusos em imagens contendo cenas de ruas [Li, Gu e Kanade, 2011] e estimação das dimensões de veículos em cenas de vídeos estereoscópicos [Ratajczak et al., 2013].

A detecção de veículos utilizando modelos pontuais divide-se em duas partes, a detecção da área do veículo e a localização de partes específicas do veículo como rodas, placas, etc. Neste trabalho, o enfoque está na segunda parte onde será realizada a modelagem de pontos característicos do objeto de estudo utilizando técnicas de modelagem da forma como o PDM, através da localização da forma, vários tipos de problemas

podem ser solucionados automaticamente como cálculo das dimensões do veículo, localização da placa, identificação do modelo do veículo, etc.

1.1 Motivação

Por muitos anos, os pesquisadores têm investigado Sistemas de Transporte Inteligente, do inglês *Intelligent Transport Systems* — ITS, planejamento de transporte e aplicações em engenharia de tráfego para extrair informações úteis e precisas através da análise de imagens. Tais informações em alguns casos permitem realizar o controle de fluxo, contagem de veículos, rastreamento, medição de velocidade, reconhecimento de placas, etc [Hadi, Sulong e George, 2014].

Para Zang et al. [2008], os sistemas tradicionais de detecção de veículos podem estar em decadência por não reconhecer bem os veículos. Isso acontece devido a ruídos e artefatos existentes na imagem, causados por condições indesejadas como sombras e oclusões. O bom funcionamento dos ITS depende da utilização de técnicas robustas de análise de imagens.

Vários métodos para detecção de carros foram sugeridos pela literatura, por exemplo, Acharya, Lee e Chen [2012] criaram um sistema para dispositivos móveis que detecta veículos em tempo real, o sistema foi implementado usando uma variação do *Adaboost* [Freund e Schapire, 1995]. Madhogaria et al. [2015] criaram uma combinação de algoritmos para detectar veículos de imagens aéreas, a abordagem consiste em extrair características baseadas em Histograma de Gradientes Orientados, do inglês *Histogram of Oriented Gradients* — HOG [Dalal e Triggs, 2005] de *patches*¹ e classificar com Máquinas de Vetores de Suporte, do inglês *Support Vector Machine* — SVM, por último, um modelo estatístico generativo é usado para gerar um *ranking* de cada *patch*.

Outros trabalhos, a exemplo de Ratajczak et al. [2013] e Zia, Stark e Schindler [2013] empregaram classificadores que avaliam as formas existentes na imagem para detectar a presença de veículos, estes trabalhos são baseados em Modelos de Formas Deformáveis. Ratajczak et al. [2013] apresentaram um novo método para estimar as dimensões dos veículos usando análise de vídeo estereoscópico baseando-se no Modelo de Aparência Ativa, do inglês *Active Appearance Model* — AAM. Zia, Stark e Schindler [2013] propuseram uma abordagem baseada em partes para detectar e modelar veículos 3D. Através da estrutura baseada em partes, o Modelo de Forma Ativa consegue restringir as variações admissíveis da forma permitindo assim reconhecer objetos com partes ocluídas.

Segundo Li, Gu e Kanade [2011], o modelo *Active Shape Model* é um dos modelos deformáveis de maior sucesso no alinhamento de objetos como órgãos e outras formas ativas como faces. Wimmer, Soza e Hornegger [2009] utilizaram ASM com sucesso na segmentação de órgãos, Ma e Ren [2011] aplicaram-no no rastreamento do corpo humano e Liu e Lovell [2005] na extração de gestos da mão.

¹*patch* consiste em um recorte da imagem, neste trabalho foi usado como amostra para treinamento.

Assim como no trabalho de Li, Gu e Kanade [2011] que empregou ASM na localização de veículos em diferentes posições, este também se propõe a investigar o emprego de ASM no alinhamento da forma dos automóveis. Ao contrário do primeiro que construiu uma máquina empregando *Random Forest* e ASM, neste a detecção do veículo é feita por ASM e Máquinas de Vetores de Suporte. Classificadores SVM são usados para identificar pontos de controle (*landmarks*) em uma vizinhança de 10×10 *pixels*. Por outro lado, admite-se que veículos sejam bem caracterizados por sua forma, o que viabiliza a utilização de um modelo de forma ativa. Por estas razões optou-se em estratificar a tarefa em dois níveis, com as escolhas de classificadores que potencialmente são mais promissores em cada um deles.

1.2 Objetivos

Partindo da hipótese de que o sucesso da detecção de veículos está relacionado com a utilização de técnicas robustas de processamento de imagens, esta dissertação tem como objetivo principal propor uma abordagem para localizar a forma dos veículos em imagens através da combinação de técnicas como modelagem pontual e classificação. A abordagem proposta é dividida em dois níveis. O primeiro nível realiza a detecção de *landmarks*, ou pontos de controle, através de máquinas de vetores de suporte. No segundo nível, a forma é ajustada aos pontos identificados pelo primeiro nível usando o modelo de forma ativa. Entre os objetivos específicos, cabe mencionar os seguintes:

- Realização de um levantamento bibliográfico sobre detecção de veículos em imagens que utilizam modelos pontuais;
- Construção de um detetor de veículo unindo modelagem da forma com classificadores de pontos de controle;
- Realização de testes com a abordagem proposta para verificar sua eficiência;
- Promoção de discussão referente à classificação automática dos pontos de controle e detecção da forma do veículo em imagens.

1.3 Metodologia

Quanto à natureza da pesquisa, este estudo se caracteriza como uma pesquisa explicativa, pois visa investigar métodos que possam contribuir com o melhoramento dos sistemas de detecção de automóveis. “Este é o tipo de pesquisa que mais aprofunda o conhecimento da realidade, porque explica a razão, o porquê das coisas” [Gill, 2008].

Inicialmente foi realizada uma pesquisa sobre o estado atual da detecção de veículos e um estudo sobre o referencial teórico acerca da abordagem a ser desenvolvida, como:

modelos deformáveis, extração de características e aprendizado de máquina. A Figura 1.1 esquematiza o processo utilizado na construção da abordagem:

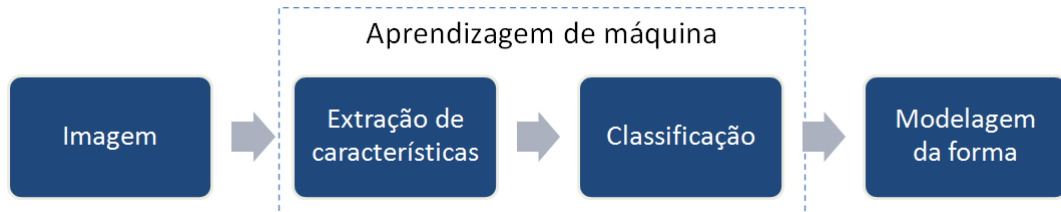


Figura 1.1: Visão geral da abordagem

A implementação da abordagem divide-se em três etapas:

- A primeira etapa é a extração de características, histogramas direcionais, capturados de *patches* de dimensões 40×40 *pixels* — *patches*, ou seções de imagens de baixa dimensão. Neste trabalho foi usada uma vizinhança de 40×40 *pixels* em razão deste se apoiar, em parte, no trabalho de [Li, Gu e Kanade, 2011], onde também foi usada a mesma abordagem;
- As características obtidas servem de entrada para o treinamento do classificador. Existem diferentes classificadores SVM para diferentes pontos de controle;
- Os resultados obtidos pela classificação são utilizados pelos modelos de formas na busca pelos pontos de controle.

Por fim, os algoritmos serão avaliados e analisados a fim de verificar a eficiência da abordagem desenvolvida.

1.4 Estrutura da Dissertação

Essa dissertação está organizada em seis capítulos, a seguir o conteúdo de cada capítulo é apresentado de forma resumida:

- Capítulo 2: Descreve os conceitos necessários para a construção dos modelos pontuais e seu variante o ASM. A construção do Modelo Pontual de Distribuição é dividida em três etapas: etiquetagem, alinhamento e estudo das variações admissíveis, já a construção do Modelo de Forma Ativa divide-se em quatro: obtenção da estimativa inicial, cálculo dos deslocamentos, obtenção dos parâmetros de ajuste da forma e por fim, atualização destes parâmetros. Todas essas etapas estão descritas de forma detalhada neste capítulo.

- Capítulo 3: Neste capítulo são descritos alguns trabalhos relacionados a aplicação dos modelos pontuais em imagens ou vídeos de veículos, algumas abordagens utilizam a versão original dos modelos pontuais, no caso do ASM, a busca pelo objeto se baseia em informações da forma e dos níveis de cinza, outras abordagens incorporaram outros métodos aos modelos pontuais para tornar os modelos ainda mais eficientes, como em Li, Gu e Kanade [2011], que serviu de base para a construção da abordagem utilizada nesta dissertação.
- Capítulo 4: Descreve a metodologia utilizada na construção da abordagem para alinhamento da forma em imagens de veículos.
- Capítulo 5: Neste capítulo, são apresentados os experimentos realizados, além dos resultados e a análise de cada um deles, por fim é feita uma comparação entre a abordagem desenvolvida e a abordagem que serviu de inspiração para a realização deste.
- Capítulo 6: Por último, são apresentadas as considerações finais bem como sugestões de trabalhos futuros.

Capítulo 2

Modelos Pontuais

Os modelos pontuais de distribuição permitem organizar informações estatísticas de um conjunto de pontos, como a forma de um rosto, do contorno da mão ou de um órgão. A partir de um conjunto de imagens do objeto de estudo são obtidas coordenadas de pontos do contorno que melhor representam a forma. Para obter conhecimentos estatísticos deste conjunto de treinamento são realizadas sucessivas operações de rotação, escalonamento e translação até que todas as amostras estejam alinhadas. O modelo pontual é obtido através do estudo das variações admissíveis do conjunto de pontos.

O modelo pontual de distribuição serve de base para a construção dos seus variantes, o modelo pontual nada mais é que a forma média e as variações admissíveis de um conjunto de pontos do objeto, além disso, fornece informações sobre os níveis de cinza na região de cada ponto. Um dos principais variantes dos modelos pontuais é o Modelo de Forma Ativa (ASM), o ASM utiliza as informações do PDM na busca por regiões da imagem que mais se assemelham com o modelo construído, as instâncias do modelo são deformadas de acordo com as variações admissíveis, assim, evita-se que a forma se mova arbitrariamente perdendo sua forma geométrica.

Neste capítulo são estudados os passos necessários para a construção do Modelo Pontual de Distribuição e do Modelo de Forma Ativa. Existem diferenças entre os modelos originais e os modelos construídos neste trabalho, a principal diferença é que no primeiro são utilizadas informações dos níveis de cinza enquanto o segundo utiliza métodos de aprendizagem de máquina para realizar a busca em novas imagens.

2.1 Modelo Pontual de Distribuição (PDM)

A construção do Modelo Pontual de Distribuição, como sugerido em Cootes et al. [1992] possui as etapas apresentadas a seguir:

2.1.1 Etiquetagem

Segundo Cootes et al. [1992], os pontos de controle devem ser capturados na mesma posição do objeto em todas as imagens, se houver um erro grosseiro na coleta, por exemplo, o ponto for coletado em diferentes regiões, todo o processo de modelagem da forma irá falhar.

De acordo com Cootes et al. [1995], bordas ou arestas com orientação particular ou pontos de curvatura extrema são preferidos, pois podem ser identificados mais facilmente por um processo automático. Esses pontos podem ser coletados manualmente por se tratar de um processo simples, no entanto, é uma tarefa cansativa, pois os pontos devem ser colocados corretamente em todas as imagens do conjunto de treinamento.

Embora existam trabalhos que realizem a etapa de etiquetagem de forma automática ou semiautomática como em Hill e Taylor [1994], Baumberg e Hogg [1994] e Brett, Hill e Taylor [1997], neste, a etiquetagem foi realizada de forma manual em imagens contendo veículos em posição frontal. Foram coletadas coordenadas de seis pontos de controle para a construção do modelo pontual onde quatro deles estão localizados nos cantos do para-brisa (Ponto 1, Ponto 2, Ponto 3 e Ponto 6) e dois no contorno do farol (Ponto 4 e Ponto 5), Figura 2.1. Esses pontos foram escolhidos com base nos pontos usados em Li, Gu e Kanade [2011], por estar em torno do contorno do veículo e por possuírem características semelhantes aos critérios adotados em Cootes et al. [1995], os pontos foram obtidos manualmente.



Figura 2.1: Pontos de controle

2.1.2 Alinhamento (Análise Generalizada de Procrustes)

Os pontos de controle devem ser alinhados para que seja possível fazer a identificação de seu arranjo topológico. Cootes et al. [1992] sugerem a utilização de uma variação

do algoritmo de Análise Generalizada de Procrustes desenvolvido por Gower [1975], cujo objetivo é alinhar os pontos correspondentes através de sucessivas operações de translação, rotação e escalonamento. A Análise Generalizada de Procrustes pode ser descrito na forma de um algoritmo:

- O primeiro passo consiste em rotacionar, escalonar e transladar cada amostra do conjunto de pontos com a primeira amostra desse conjunto;

Em seguida, os seguintes passos devem ser repetido sucessivas vezes até que um número de iterações seja atingido ou até não existir mudanças significativas:

- Obter a média do conjunto alinhado;
- Rotacionar, escalonar e transladar a média para que se alinhe com a primeira amostra do conjunto;
- Rotacionar, escalonar e transladar cada amostra do conjunto de pontos para que se alinhe com a média calculada;

A seguir serão descritos os passos para se obter a forma média e as operações de translação, escalonamento e rotação. Considere que cada instância \mathbf{x}_i de um objeto seja representada pela seguinte sequência de vértices do plano cartesiano, como a matriz na Equação 2.1.

$$\mathbf{x}_i \begin{bmatrix} x_{i0} & y_{i0} \\ \vdots & \vdots \\ x_{ik} & y_{ik} \\ \vdots & \vdots \\ x_{in-1} & y_{in-1} \end{bmatrix} \quad (2.1)$$

Seja $\mathbf{x}_j = \begin{bmatrix} x_{j0} & y_{j0} \\ \vdots & \vdots \\ x_{jn-1} & y_{jn-1} \end{bmatrix}$ a forma média a qual será usada como referência para fazer o alinhamento dos pontos.

Na primeira etapa é aplicada a translação para mover a média do conjunto para a origem através da expressão:

$$\mathbf{x}_i \rightarrow \mathbf{x}_i - \bar{\mathbf{x}}, \quad (2.2)$$

onde,

$$\bar{\mathbf{x}} = \frac{1}{N} \sum_{i=1}^N \mathbf{x}_i, \quad (2.3)$$

onde N representa a quantidade de amostras de treinamento. Após aplicar a translação calcula-se o fator de escala s_i , que é dado pela Equação 2.4,

$$s_i = \sqrt{\frac{n}{\sum_{k=1}^n \mathbf{x}_{ik}^2}} \quad (2.4)$$

cada instância, \mathbf{x}_i , é atualizada como:

$$\mathbf{x}_i \rightarrow \mathbf{x}_i * s_i, \quad (2.5)$$

o último parâmetro a ser encontrado é o ângulo de rotação θ_i :

$$\theta_i = \arctan \frac{\text{sen} \left(\begin{bmatrix} x_{i0} \\ \vdots \\ x_{in-1} \end{bmatrix}^t \begin{bmatrix} y_{j0} \\ \vdots \\ y_{jn-1} \end{bmatrix} - \begin{bmatrix} y_{i0} \\ \vdots \\ y_{in-1} \end{bmatrix}^t \begin{bmatrix} x_{j0} \\ \vdots \\ x_{jn-1} \end{bmatrix} \right)}{\text{cos} \left(\begin{bmatrix} x_{i0} \\ \vdots \\ x_{in-1} \end{bmatrix}^t \begin{bmatrix} y_{j0} \\ \vdots \\ y_{jn-1} \end{bmatrix} + \begin{bmatrix} y_{i0} \\ \vdots \\ y_{in-1} \end{bmatrix}^t \begin{bmatrix} x_{j0} \\ \vdots \\ x_{jn-1} \end{bmatrix} \right)} \quad (2.6)$$

\mathbf{x}_j é a forma de referência para a qual todas as outras formas se alinham. Por fim, cada forma é ajustada a \mathbf{x}_j através da rotação a seguir:

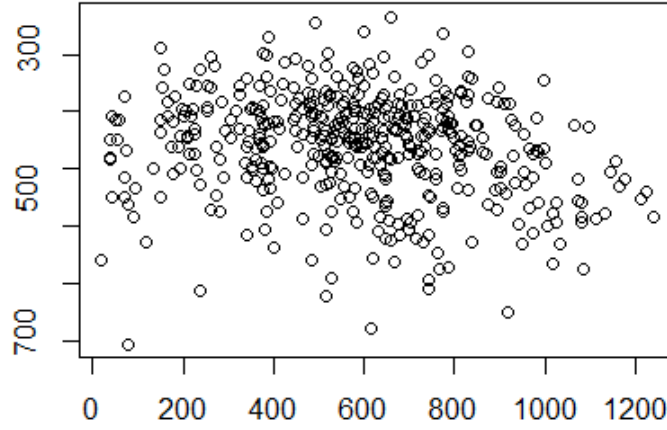
$$\mathbf{x}_i \rightarrow \mathbf{x}_i * \begin{bmatrix} \cos \theta & \text{sen} \theta \\ -\text{sen} \theta & \cos \theta \end{bmatrix} \quad (2.7)$$

Estas operações devem ser realizadas repetidas vezes para que todas as formas do conjunto de treinamento sejam alinhadas. Segundo Cootes et al. [1992] a condição de convergência pode ser verificada através da diferença média entre as transformações necessárias para alinhar cada forma com a forma média recalculada. Isto é, caso a transformação da forma seja próxima à identidade, o critério de convergência terá sido alcançado.

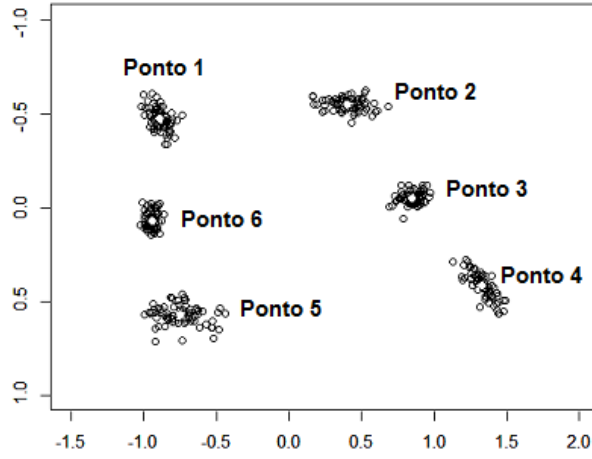
Os pontos desalinhados, obtidos em diferentes escalas e posições, Figura 2.2(a), depois de submetidos às sucessivas operações de escalonamento, translação e rotação são alinhados em posições próximas, como apresentado na Figura 2.2(b).

2.1.3 Variações Admissíveis

Depois de obter as instâncias de treinamento e alinhá-las com a Análise Generalizada de Procrustes, o próximo passo é extrair informações como média e variância das formas para finalmente obter o modelo pontual da forma estudada. Considere que cada instância \mathbf{x}_i de um objeto seja representada por uma sequência de vértices do plano cartesiano, expressa como uma matriz como na Equação (2.1).



(a) Amostras desalinhadas



(b) Amostras alinhadas

Figura 2.2: Alinhamento de amostras através do algoritmo Análise Generalizada de Procrustes; (a) amostras desalinhadas, (b) amostras alinhadas

A forma média $\bar{\mathbf{x}}$ é calculada como na Equação 2.3 e a matriz de covariância \mathbf{S} é obtida por:

$$\mathbf{S} = \frac{1}{N} \sum_{i=1}^N (\mathbf{x}_i - \bar{\mathbf{x}})(\mathbf{x}_i - \bar{\mathbf{x}})^t, \quad (2.8)$$

em seguida calculam-se os intervalos de variação para os pontos através dos autovalores λ_i e autovetores \mathbf{p}_i da matriz \mathbf{S} . Assim, as variações mais significativas podem ser representadas pelos t autovetores correspondentes aos t maiores autovalores. As instâncias de uma forma podem ser representadas por:

$$\mathbf{x} = \bar{\mathbf{x}} + \mathbf{P}\mathbf{b}, \quad (2.9)$$

onde $\mathbf{P} = (\mathbf{p}_1 \mathbf{p}_2 \cdots \mathbf{p}_t)$ é a matriz dos primeiros autovetores e $\mathbf{b} = (b_1 b_2 \cdots b_t)^t$ é o vetor de pesos dos autovalores.

As equações apresentadas anteriormente permitem a generalização de novos exemplos da forma variando o parâmetro b_i dentro de limites aceitáveis. Admitindo-se que a variação aceitável se concentra entre três desvios padrões em torno da média, o limite máximo e mínimo que a forma pode assumir é delimitado pelo intervalo na Equação 2.10:

$$-3\sqrt{\lambda_i} \leq b_i \leq 3\sqrt{\lambda_i}. \quad (2.10)$$

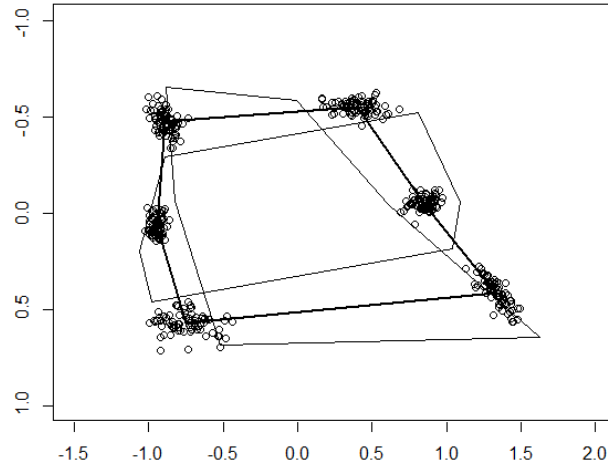


Figura 2.3: Resultado da Análise de Componentes Principais

O resultado da Análise de Componentes Principais é apresentado na Figura 2.3. A linha em destaque representa a forma média dos pontos. Logo, a forma deve ser deformada entre o limite inferior e superior para que assuma uma forma aceitável.

2.1.4 Perfis de Intensidade

Uma vez calculadas as variações admissíveis para os pontos de controle, o próximo passo envolve extrair as características de variação de intensidade da região próxima ao ponto. O cálculo para obter uma determinada quantidade de *pixels* que represente o ponto de controle é feito da seguinte forma:

A partir de um ponto j da imagem i de treinamento, são amostrados os níveis de cinza n_{acima} *pixels* acima e n_{abaixo} *pixels* abaixo do ponto de controle, assim o vetor de intensidades \mathbf{g}_{ij} possui dimensão $n_p = n_{acima} + n_{abaixo} + 1$ e é dado pela Equação 2.11. Após obter os perfis de cinza de todos os pontos e imagens de treinamento é possível criar um modelo estatístico das intensidades.

$$\mathbf{g}_{ij} = [g_{ij0} \quad g_{ij1} \quad \cdots \quad g_{ij(n_p-1)}]^t \quad (2.11)$$

O perfil derivativo tem dimensão $n_p - 1$ dado pela equação:

$$\mathbf{g}'_{ij} = [g_{ij1} - g_{ij0} \quad g_{ij2} - g_{ij1} \quad \cdots \quad g_{ij(n_p-1)} - g_{ij(n_p-2)}]^t, \quad (2.12)$$

a normalização da derivada do perfil é dada por:

$$\mathbf{g}'_{ij} = \frac{\mathbf{g}'_{ij}}{\sum_{k=0}^{n_p-2} \mathbf{g}'_{ijk}} \quad (2.13)$$

assim, é possível calcular a média $\bar{\mathbf{g}}_j$ e a matriz de covariância \mathbf{C}_{gj} para cada ponto em todas as amostras de treinamento N :

$$\bar{\mathbf{g}}_j = \frac{1}{N} \sum_{i=1}^N \mathbf{g}'_{ij}, \quad (2.14)$$

$$\mathbf{C}_{gj} = \frac{1}{N} \sum_{i=1}^N (\mathbf{g}'_{ij} - \bar{\mathbf{g}}_j)(\mathbf{g}'_{ij} - \bar{\mathbf{g}}_j)^t. \quad (2.15)$$

No algoritmo ASM, a busca pela forma em uma imagem de teste envolve não só informações geométricas como também perfis de intensidade dos *pixels* próximos aos pontos. Considerando que os níveis de cinza em uma mesma região de diferentes imagens são semelhantes, desta forma, essa informação é utilizada na fase de busca para encontrar a região mais parecida com o perfil obtido.

2.2 Busca da Forma na Imagem (Modelo de Forma Ativa)

O Modelo de Forma Ativa é um caso particular do Modelo Pontual de Distribuição empregado para realizar busca da forma do objeto modelado em imagens. Uma forma inicial é deformada iterativamente para encontrar a maior semelhança entre o objeto modelado e os dados da imagem, no ASM são realizadas quatro etapas para encontrar a forma em uma imagem: estimar a forma inicial, calcular um deslocamento em cada ponto para se ajustar a imagem, encontrar os parâmetros para modificar a forma e pose e ajustar os parâmetros.

2.2.1 Estimativa Inicial

De acordo com Vasconcelos, Tavares e Ribeiro [2005] o posicionamento inicial do modelo é um fator determinante para o sucesso das etapas posteriores. A estimativa inicial \mathbf{x}_i é obtida através da rotação θ , escalonamento s e translação \mathbf{t} da forma \mathbf{x}_1 :

$$\mathbf{x}_i = \mathbf{M}(s_i, \theta_i)[\mathbf{x}_1] + \mathbf{t}_i, \quad (2.16)$$

onde

$$\mathbf{M}(s_i, \theta_i) = s_i \begin{bmatrix} \cos \theta_i & -\sin \theta_i \\ \sin \theta_i & \cos \theta_i \end{bmatrix}, \quad (2.17)$$

$$\mathbf{t}_i = [t_{xi} \quad t_{yi} \quad t_{xi} \quad \cdots \quad t_{xi} \quad t_{yi}]^t \quad (2.18)$$

e, $\mathbf{x}_1 = \bar{\mathbf{x}} + d\mathbf{x}_1$, com $d\mathbf{x}_1 = \mathbf{P}\mathbf{b}_1$.

Assim, a estimativa inicial \mathbf{x}_i pode ser escrita na forma:

$$\mathbf{x}_i = \mathbf{M}(s_i, \theta_i)[\bar{\mathbf{x}} + d\mathbf{x}_i] + \mathbf{t}_i, \quad (2.19)$$

2.2.2 Cálculo do Movimento Sugerido para Cada Ponto do Modelo

Após obter a estimativa inicial inicia-se a busca pela forma, o cálculo dos deslocamentos consiste em um estudo da região em torno de cada ponto para encontrar os ajustes que irão movê-los para as melhores posições na vizinhança. O cálculo é feito com base nas informações estatísticas dos perfis de níveis de cinza.

Deve ser obtido um perfil com dimensão n_s centralizado no ponto de controle i , a dimensão n_s deve ser maior que a dimensão n_p do perfil obtido no treinamento do modelo. O perfil de tamanho n_s é dividido em subperfis de tamanho n_p , assim é possível comparar os perfis de busca com o perfil de treinamento através de uma função de semelhança, como por exemplo, a distância de *Mahalanobis* [Cootes, Taylor e Lanitis, 1994], Figura 2.4. Deste modo, a melhor posição dentro da região de busca é selecionada para a qual o ponto i deve se mover tornando-se viável o cálculo do deslocamento $d\mathbf{x}_i$.



Figura 2.4: Busca por perfis de níveis de cinza

2.2.3 Computação das Alterações dos Parâmetros de Pose e Forma

Depois de encontrar o deslocamento associado a cada ponto é necessário ajustar a estimativa ao deslocamento calculando os parâmetros da pose (rotação, escalonamento e translação) e da forma (os pesos dos componentes principais \mathbf{b}), assim torna-se possível aproximar a estimativa \mathbf{x}_i ao deslocamento $\mathbf{x}_i + d\mathbf{x}_i$ sem ultrapassar os limites aceitáveis da forma [Cootes et al., 1995]. Neste trabalho, utilizou-se a Análise Generalizada de Procrustes para obter a rotação $d\theta$, escalonamento ds e translação dt objetivando o alinhamento da estimativa \mathbf{x}_i com o deslocamento $d\mathbf{x}_i$.

Conhecendo os valores de $(1+ds)$, $d\theta$ e dt , o cálculo do deslocamento é obtido através da equação:

$$\mathbf{M}(s_i(1+ds), \theta_i + d\theta)[\mathbf{x}_1 + d\mathbf{x}] + t_i + dt = \mathbf{x}_i + d\mathbf{x}_i. \quad (2.20)$$

Resolvendo a Equação 2.20 obtém-se a Equação 2.21:

$$d\mathbf{x} = \mathbf{M}(s_i(1+ds)^{-1}, -(\theta_i + d\theta))[\mathbf{M}(s_i, \theta_i)[\mathbf{x}_1] + d\mathbf{x}_i - dt] - \mathbf{x}. \quad (2.21)$$

A Equação 2.21 é uma forma de calcular os deslocamentos da estimativa \mathbf{x}_i , porém esse alinhamento em geral pode tornar a forma inconsistente, logo, é necessário aplicar restrições para que o deslocamento da estimativa sempre assuma uma forma aceitável. Para aplicar as restrições da forma, $d\mathbf{x}'$ é projeção de $d\mathbf{x}$ em um espaço de t dimensões descritas pelos primeiros eixos principais (ou seja, o espaço t de \mathbf{P} é definido pelo vetores das componentes principais). Assim, o cálculo do deslocamento torna-se uma aproximação entre a estimativa e o deslocamento, matematicamente tem-se:

$$d\mathbf{x}' = \mathbf{A}d\mathbf{x}, \quad (2.22)$$

onde \mathbf{A} é matriz de projeção dada por $\mathbf{A} = \mathbf{P}(\mathbf{P}^t\mathbf{P})^{-1}\mathbf{P}^t$.

As colunas de \mathbf{P} são ortogonais e \mathbf{P} não é uma matriz quadrada, tem-se $\mathbf{P}^t\mathbf{P} = \mathbf{I}$, e assim $d\mathbf{x}' = \mathbf{P}\mathbf{P}^td\mathbf{x}$.

Assim, o deslocamento $\mathbf{x}_1 + d\mathbf{x}$ passa a ser $\mathbf{x}_1 + d\mathbf{x}'$. Representando $d\mathbf{x}'$ como $d\mathbf{x}' = \mathbf{P}d\mathbf{b}'$, multiplicando por \mathbf{P}^t , é obtido:

$$d\mathbf{b}' = \mathbf{P}^td\mathbf{x}'. \quad (2.23)$$

2.2.4 Atualização dos Parâmetros

Depois de obter os parâmetros da pose e da forma uma nova estimativa $\mathbf{x}_i^{(1)}$ deve ser obtida através da equação:

$$\mathbf{x}_i^{(1)} = \mathbf{M}(s_i(1+ds), \theta_i + d\theta)[\mathbf{x}_1 + \mathbf{P}d\mathbf{b}'] + t_i + dt. \quad (2.24)$$

Os parâmetros são atualizados e todo o processo é reiniciado sendo que $\mathbf{x}_i^{(1)}$ é considerada a nova estimativa inicial e $\mathbf{x}_i^{(2)}$ representa a estimativa a ser calculada. Os parâmetros são atualizados como:

$$t_{x_i} \rightarrow t_{x_i} + dt_x \quad (2.25)$$

$$t_{y_i} \rightarrow t_{y_i} + dt_y \quad (2.26)$$

$$s_i \rightarrow s_i(1 + ds) \quad (2.27)$$

$$\theta_i \rightarrow \theta_i + d\theta \quad (2.28)$$

$$\mathbf{b} \rightarrow \mathbf{b} + d\mathbf{b}' \quad (2.29)$$

esse processo é repetido até que as mudanças entre a nova estimativa e a anterior sejam mínimas, ou seja, até que o critério de convergência no alinhamento da estimativa tenha sido atingido, ou até que um dado número de iterações seja completado.

Capítulo 3

Modelos Pontuais e Detecção de Veículos

Nesta seção são apresentados alguns trabalhos relacionados ao modelo ASM, destacando os métodos utilizados no processo de extração de características de cada abordagem.

No ASM, proposto por Cootes e Taylor [1993], a fase de busca por novas instâncias do objeto consiste no estudo da região em torno dos pontos de controle sendo utilizadas características 1-D baseadas nos perfis de intensidade dos *pixels*. Esses perfis são comparados pela função de *Mahalanobis*, a fim de obter o perfil com maior semelhança ao de treinamento para que a forma se desloque em direção ao ponto de controle, como ilustrado na Figura 3.1.

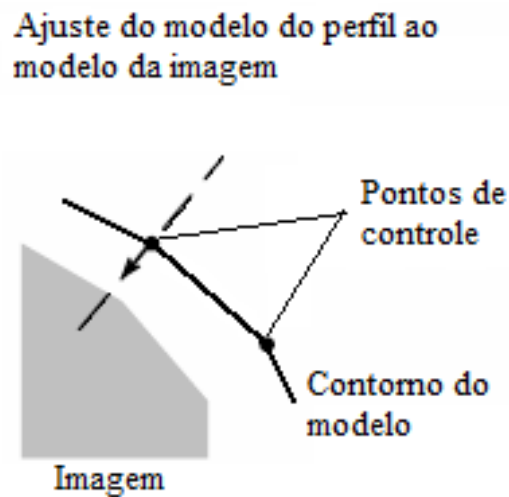


Figura 3.1: Deslocamento associado ao ponto de controle
Fonte: Adaptado de Cootes et al. [1995]

Ratajczak et al. [2013] apresentaram uma proposta para estimar as dimensões de veículos baseada em Análise de Vídeo Estereoscópico. Foram utilizadas duas câmeras que capturam simultaneamente a imagem do veículo e obtém-se a projeção da imagem em 3D, Figura 3.2. Para localizar os pontos de controle em ambas as imagens Ratajczak et al. [2013] optaram em utilizar o Modelo de Aparência Ativa, pois este modelo consegue determinar a localização da forma do veículo com precisão. A forma consiste em um conjunto de pontos que são agrupados em múltiplos polígonos, a aparência ou textura do modelo consiste em todos os *pixels* que estão dentro dos polígonos da forma definida.

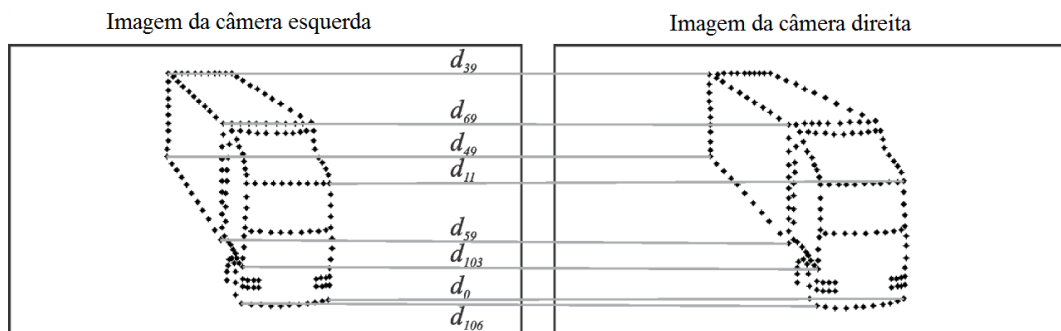


Figura 3.2: Modelagem com AAM

Fonte: Ratajczak et al. [2013]

Tsai et al. [2012] usaram modelos de veículos 3D mais delicados que fornecem restrições ricas para combinar veículos de forma confiável e refinar o alinhamento 3D para 2D. A abordagem proposta baseia-se na criação do Modelo de Forma Ativa para rastrear os veículos. A busca pelas correspondências corretas entre os pontos projetados e os pontos do objeto é realizada através do algoritmo *Point Registration* — PR enquanto o método de mínimos quadrados é utilizado para otimizar a forma e os parâmetros da pose iterativamente. Tsai et al. [2012] detectaram veículos com diferentes visões em ambientes não controlados onde a pesquisa pelo veículo é baseada em partes como rodas, farol e motor. Para obter estas partes de interesse foi construído um modelo 3D utilizando o Modelo de Forma Ativa. Primeiramente, o modelo do veículo é projetado na imagem através da estimativa inicial onde os pontos são projetados nas bordas, em seguida, para cada ponto amostrado na borda, os pontos próximos à direção do contorno são encontrados como um conjunto de pontos candidatos. Assim como no trabalho Tsin, Genc e Ramesh [2009], este também usou a abordagem PR para encontrar uma transformação rígida ou não rígida aplicada ao conjunto de pontos projetados. Por fim, o problema de ajuste do modelo é formulado como um problema de mínimos quadrados. Esse processo é ilustrado na Figura 3.3.

Li, Gu e Kanade [2011] adotaram uma abordagem baseada em hipótese e teste randomizado que investiga o problema de alinhamento da forma na presença de *outliers*.

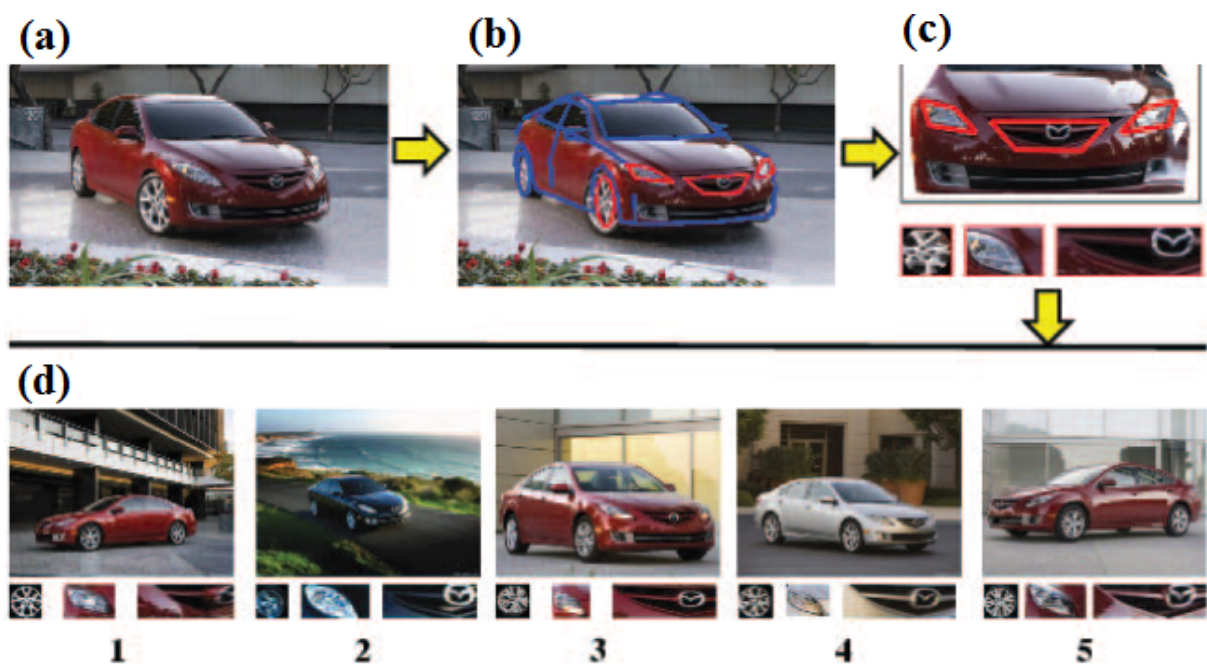


Figura 3.3: Visão geral do sistema, (a) imagem de entrada, (b) alinhamento do modelo 3D para imagem 2D, (c) Correção dos veículos para o mesmo ponto de vista e extração de partes dos veículos (por exemplo, motor, rodas, e farol) e (d) 5 melhores resultados de pesquisa através da fusão das três partes

Fonte: Tsai et al. [2012]

Uma importante contribuição deste trabalho é o algoritmo de Inferência Bayesiana da forma, a partir de uma forma parcial ou um subconjunto de pontos de controle um grande número de hipóteses é gerado por amostragem aleatória e em seguida são avaliadas para estimar os parâmetros ideais da forma e pose do objeto.

A localização de partes dos veículos em diferentes ângulos da câmera é outra contribuição do trabalho de Li, Gu e Kanade [2011]. Os exemplos de treinamento do PDM foram rotulados em cinco diferentes visões: semi frontal, perfil, semi fundo, frontal e fundo, a forma é representada por 14, 10, 14, 8 e 8 pontos de controle, respectivamente (Figura 3.4). Para o ASM foram extraídos *patches* de 40×40 *pixels* com exemplos positivos e negativos para cada ponto de controle, os descritores HOG foram computados através da sobreposição de bloco 2×2 contendo células 8×8 ao longo da imagem *patch*, os gradientes 2-D são representados por um histograma de 9 orientações. As características obtidas formaram portanto um vetor de 576 componentes.

Li, Gu e Kanade [2011] empregaram uma Floresta Aleatória (*Random Forest*) como máquina algorítmica para classificar os vetores de atributos computados na fase de extração de características. Eles empregaram 100 árvores de decisão geradas aleatoria-

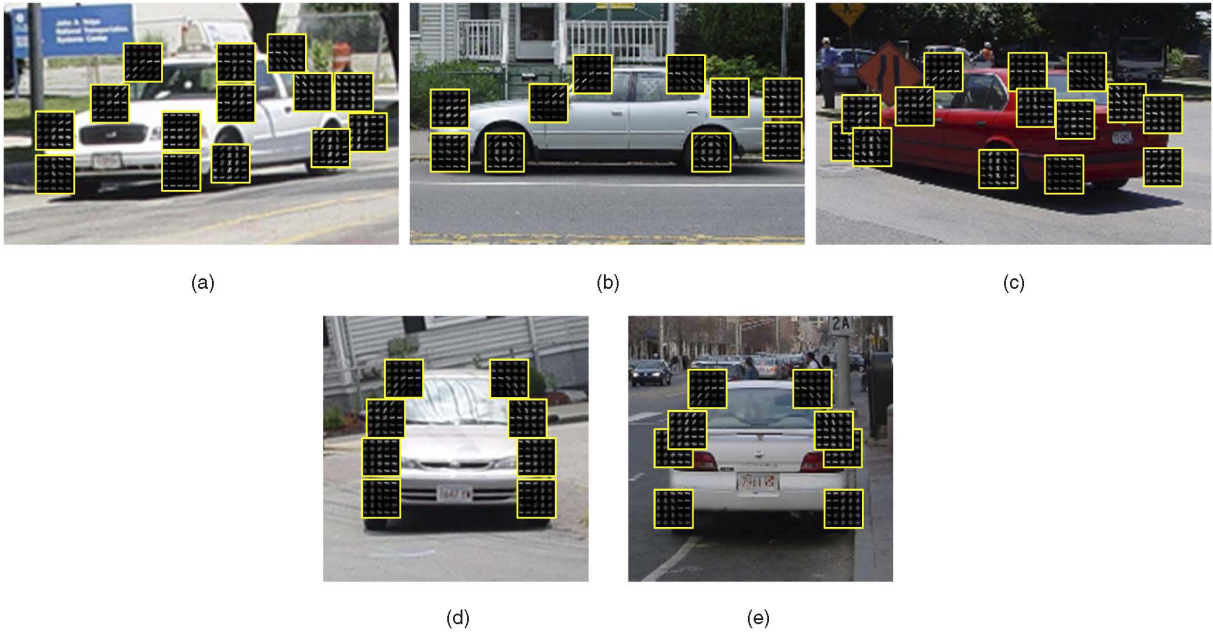


Figura 3.4: Características dos pontos de controle em diferentes visões do veículo, (a) semi frontal, (b) perfil, (c) semi fundo, (d) frontal e (e) fundo

Fonte: Li, Gu e Kanade [2011]

mente treinadas com 5.000 amostras cada uma. Todos os detetores foram combinados em uma única Floresta Aleatória.

Para a inicialização, Li, Gu e Kanade [2011] construíram diferentes detetores baseados em *boosting* para as cinco diferentes posições do veículo. Para treinar um localizador de pose, são extraídas características da forma e cor de um *patch* de entrada, o vetor de características resultante é enviado a um classificador para determinar o posicionamento do veículo. Quando as probabilidades resultantes da classificação estão acima de um limiar, a detecção do carro procede por meio de um conjunto de detetores específicos para cada vista. A saída do localizador é um retângulo e o ponto de vista correspondente, a forma média para esse ponto de vista é usada como estimativa inicial para o alinhamento ASM, este processo está ilustrado na Figura 3.5.

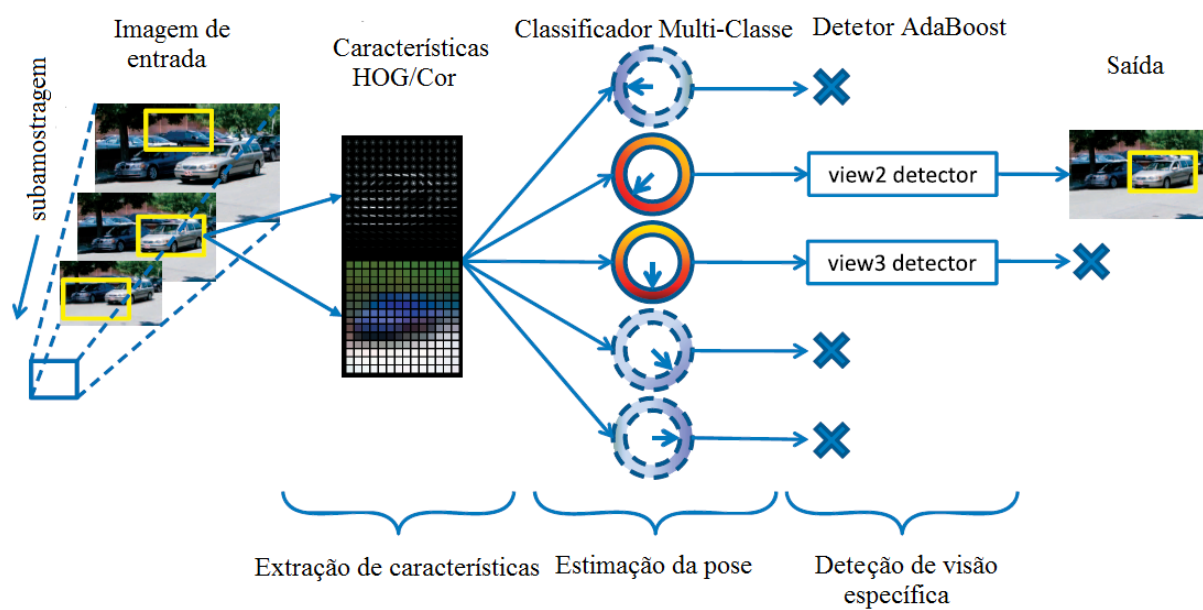


Figura 3.5: Ilustração do processo de detecção de veículos
 Fonte: Li, Gu e Kanade [2011]

Capítulo 4

Desenvolvimento da Abordagem

A abordagem proposta neste trabalho divide-se em duas etapas: construção do Modelo de Forma Ativa e do classificador de pontos de controle desenvolvido com base em características HOG e Máquinas de Vetores de Suporte. Os modelos de forma ativa e de distribuição foram implementados na linguagem R, enquanto os classificadores e os extratores de características na linguagem Python. Os classificadores SVM e Florestas Aleatórias foram construídos com a biblioteca *scikit-learn* e os extratores HOG e Padrão Binário Local, do inglês *Local Binary Pattern* — LBP com *scikit-image*, o LBP utilizado foi o circular com 24 pontos e raio igual a 3. A seguir será detalhado o modo como esta abordagem foi desenvolvida:

4.1 Base de Imagens

Foi utilizada a base de imagens *MIT Street Scene* ², que contém 3081 imagens de veículos de tamanho 1280×960 *pixels* com cenas de ruas e em reais condições de iluminação. Os veículos presentes nas cenas possuem diferentes escalas, cores e orientações. Como a base contém uma variedade ampla de imagens de veículos, optou-se em restringir as imagens para aquelas com visão frente lateral, resultando em 204 imagens.

Foi considerado o método *holdout* particionando a base de dados em 3/4 para treinamento e 1/4 para teste. Foram escolhidos seis pontos diferentes do veículo em cada imagem de treinamento para construção do modelo pontual.

Para o Modelo de Forma Ativa foram selecionados os mesmos pontos do modelo pontual sendo segmentados *patches* com amostras positivas (*patches* onde o ponto de controle está centralizado) e negativas. Os *patches* positivos são formados por pedaços da imagem de tamanho de 40×40 *pixels* onde o ponto de controle está centralizado, além disso, a vizinhança 8 de cada ponto também foi considerada como amostras positivas balanceando a quantidade de exemplos positivos e negativos. As amostras negativas

²<http://cbcl.mit.edu/software-datasets/streetscenes/>.

foram selecionadas aleatoriamente, assim, foram capturados 9 exemplos positivos e 36 negativos para cada ponto em cada imagem. A Figura 4.1 apresenta um exemplo das amostras capturadas na imagem da Figura 4.2 para o Ponto 1, onde as 9 primeiras são exemplos positivos e as outras 36 são negativas.



(a) Amostras positivas



(b) Amostras negativas

Figura 4.1: Exemplo de amostras positivas e negativas para treinamento do Modelo de Forma Ativa



Figura 4.2: Amostra de treinamento

4.2 Modelo de Forma Ativa

O Modelo de Forma Ativa é empregado para realizar a busca pela forma do objeto em imagens. Após implementar o Modelo Pontual de Distribuição como sugerido por Cootes et al. [1992] e obter as variações admissíveis (Figura 2.3) é possível realizar as buscas pelo objeto através do Modelo de Forma Ativa (ASM):

4.2.1 Estimativa Inicial

A posição inicial para cada ponto foi escolhida com auxílio de máquinas de vetores de suporte (Figura 4.3), as máquinas foram treinadas para classificar os *pixels* da imagem contendo o veículo, quanto mais alta a saída produzida pela SVM mais se aproxima do ponto de controle, assim a posição inicial foi obtida pelos seguintes passos: a forma média do modelo de distribuição pontual foi escalonada de acordo com o tamanho da imagem de teste e centralizada na imagem. Em seguida, foi selecionada uma região de 60×60 *pixels* em torno de cada ponto da forma média e dentro de cada região foi localizado o pixel mais claro, por fim, a forma média foi rotacionada, transladada e escalonada para alinhar-se com os melhores resultados obtidos através da classificação SVM das regiões (esta etapa está detalhada na seção a seguir). Esse processo resultou na estimativa inicial, como apresentado na Figura 4.4.

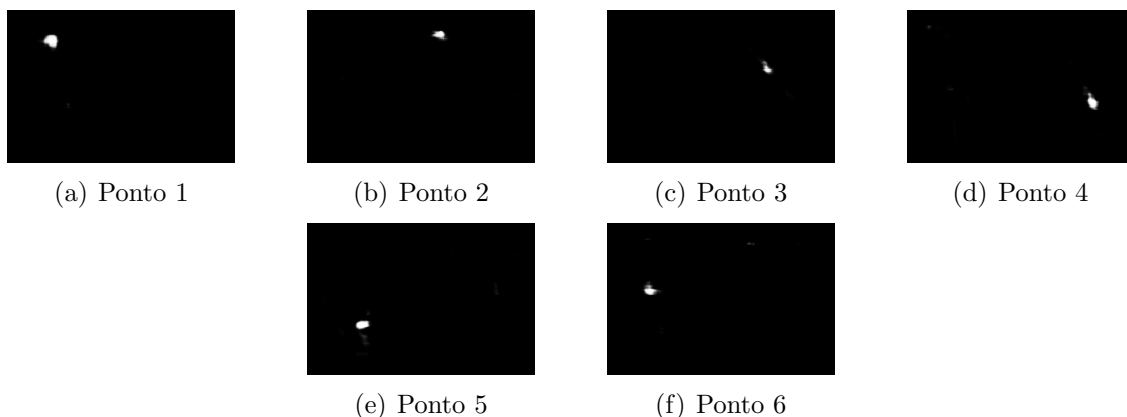


Figura 4.3: Resultados da classificação da Máquina de Vetores de Suporte; (a) – (f) Mapa de saliência obtido para cada ponto

4.2.2 Cálculo dos Deslocamentos Através do Classificador SVM

Após obter a estimativa inicial, o próximo passo é calcular os ajustes necessários para que os pontos da estimativa obtida se desloquem em direção a suas regiões equivalentes na imagem. Os ajustes são obtidos através da classificação de uma região de 10×10



Figura 4.4: Estimativa inicial

pixels em torno de cada ponto. Assim como Li, Gu e Kanade [2011], para construir o classificador, foram extraídas características dos pontos de controle através de imagem de 40×40 *pixels* onde o ponto de controle está centralizado.

Através das amostras colhidas de *patches* é criado um descritor baseado em HOG. Para computar o descritor, a imagem *patch* divide-se em blocos de 2×2 células, a célula tem um tamanho de 8×8 *pixels*, em cada célula são calculadas 9 orientações. O vetor de características obtido possui 576 posições, decorrente do deslizamento de uma janela 16×16 *pixels* na vizinhança do pixel central. Os histogramas representando as 9 orientações foram calculados com base na equações (4.1) e (4.2).

$$m(x, y) = \left((L(x+1, y) - L(x-1, y))^2 + (L(x, y+1) - L(x, y-1))^2 \right)^{1/2}, \quad (4.1)$$

$$\theta(x, y) = \tan^{-1} \left(\frac{L(x, y+1) - L(x, y-1)}{L(x+1, y) - L(x-1, y)} \right), \quad (4.2)$$

sendo que $m(x, y)$ é a magnitude e $\theta(x, y)$ é a orientação do gradiente da amostra da imagem $L(x, y)$. A Figura 4.5 mostra a média do descritor HOG para cada ponto de controle do veículo.

Existem vários métodos de aprendizagem utilizados para a detecção de objetos, como SVM, *AdaBoost* e redes neurais, [Wang et al., 2009]. Nenhum deles é garantidamente superior em todas aplicações, no entanto, há três que em geral são mais bem sucedidos, como aponta o trabalho realizado por Fernández-Delgado et al. [2014]. Neste artigo os autores experimentaram 179 algoritmos de aprendizado de máquina aplicados a 121 bases públicas do repositório UCI [Lichman, 2013]. Após todas rodadas de experimentação, verificaram que os melhores desempenhos, medido pela acurácia, foram obtidos pelos métodos *Random Forest*, *Support Vector Machine* e *Neural Network*. Por esta razão, neste trabalho foram realizados experimentos com SVM e florestas randômicas.

Originalmente, as máquinas de vetores de suporte foram empregadas para regressão (ou *Support Vector Regression* — SVR) a fim de identificar as regiões em torno dos

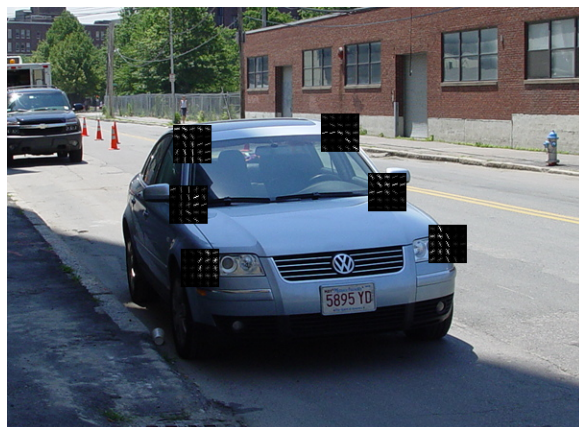


Figura 4.5: Características HOG de cada ponto

pontos de controle. Os descritores extraídos alimentam a Máquina de Vetores de Suporte para regressão, que fornece valores mais altos se a região apresentada à máquina se assemelha com o ponto que se deseja encontrar. Após o treinamento, a máquina possui a capacidade de generalizar e assim consegue classificar os exemplos não vistos.

A saída produzida pela SVR representa uma distância entre um vetor de teste e os hiperplanos canônicos no espaço de atributos. Estes valores são normalizados para o intervalo $[0, 1]$ através da adaptação do algoritmo de Platt [1999] feita em Lin, Lin e Weng [2007].

Foram realizados diversos experimentos com diferentes máquinas de regressão e diferentes métodos de extração de características, como discutido no Capítulo 5. De um modo geral, em uma perspectiva macro, em todos eles foi empregado o mesmo procedimento, cujo diagrama de blocos é apresentado na Figura 4.6.

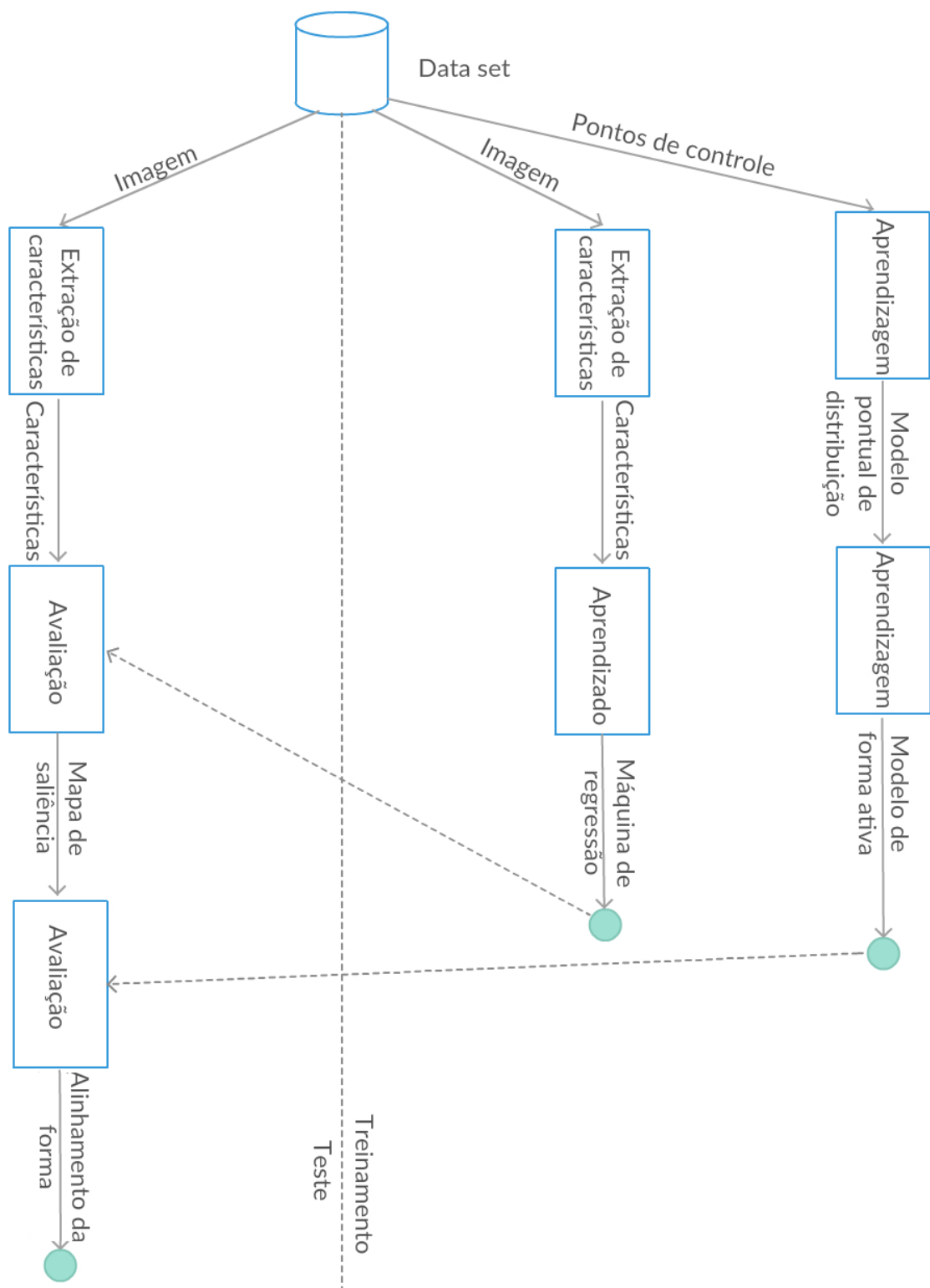


Figura 4.6: Diagrama da visão geral dos experimentos

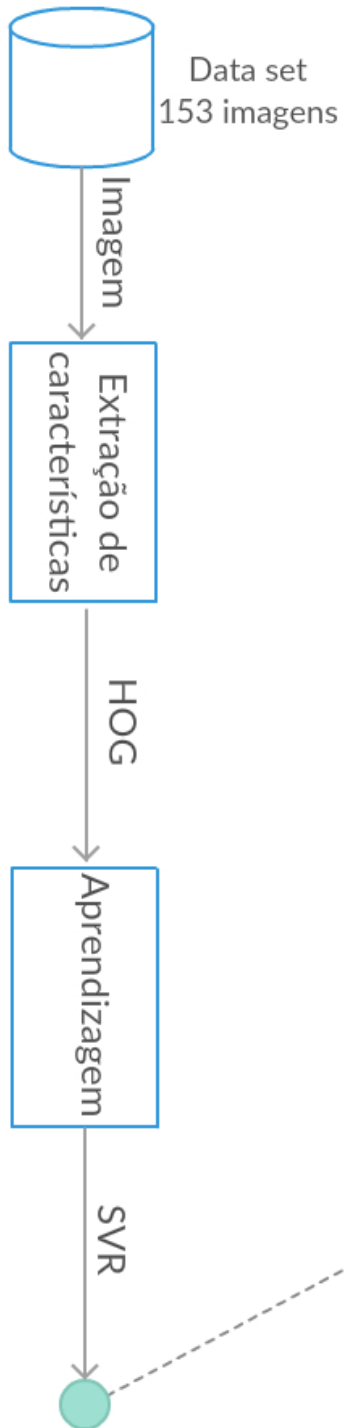
Capítulo 5

Experimentos e Resultados

Nesta seção são descritos os experimentos realizados após o desenvolvimento dos algoritmos. Foram realizados quatro, onde cada um será apresentado detalhadamente a seguir:

- Experimento I: O primeiro experimento consiste na classificação SVM das 51 imagens de teste para localizar a posição de cada ponto de controle usando características HOG, para ilustrar esse processo, o diagrama de blocos é apresentado na Figura 5.1;
- Experimento II: Neste, um segundo nível de classificação foi implementado usando máquinas SVM e características LBP para classificar 15 resultados dentre os 51 do Experimento I, Figura 5.2;
- Experimento III: As mesmas características e classificadores utilizados no Experimento I também foram usados neste, porém, os 15 resultados dentre os 51 obtidos no primeiro nível foram usados como entrada para o segundo classificador, Figura 5.3;
- Experimento IV: A combinação de florestas aleatórias e HOG foi utilizada para detetar os pontos de controle nas 51 imagens de teste, Figura 5.4.

Etapa de treinamento



Etapa de teste

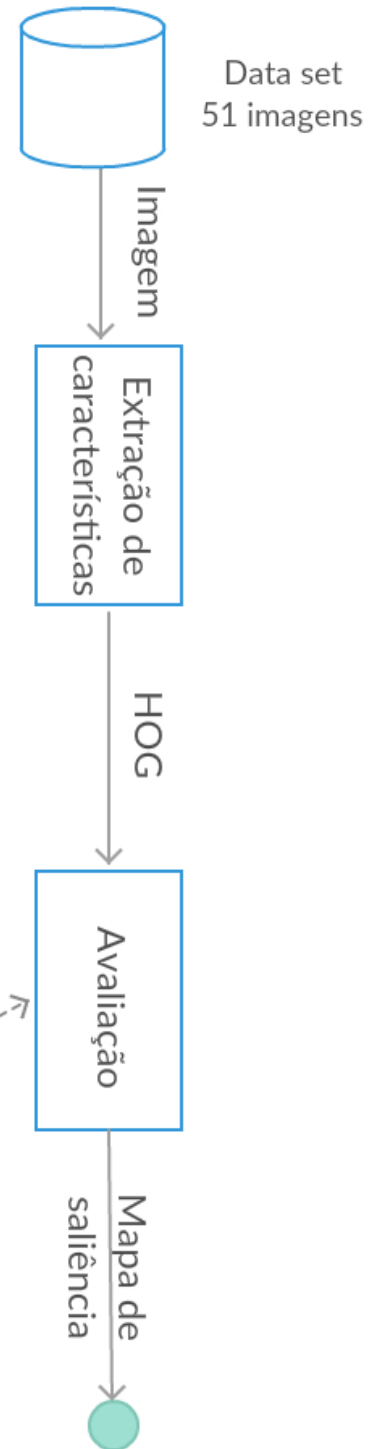


Figura 5.1: Diagrama do Experimento I

Etapa de treinamento

Etapa de teste

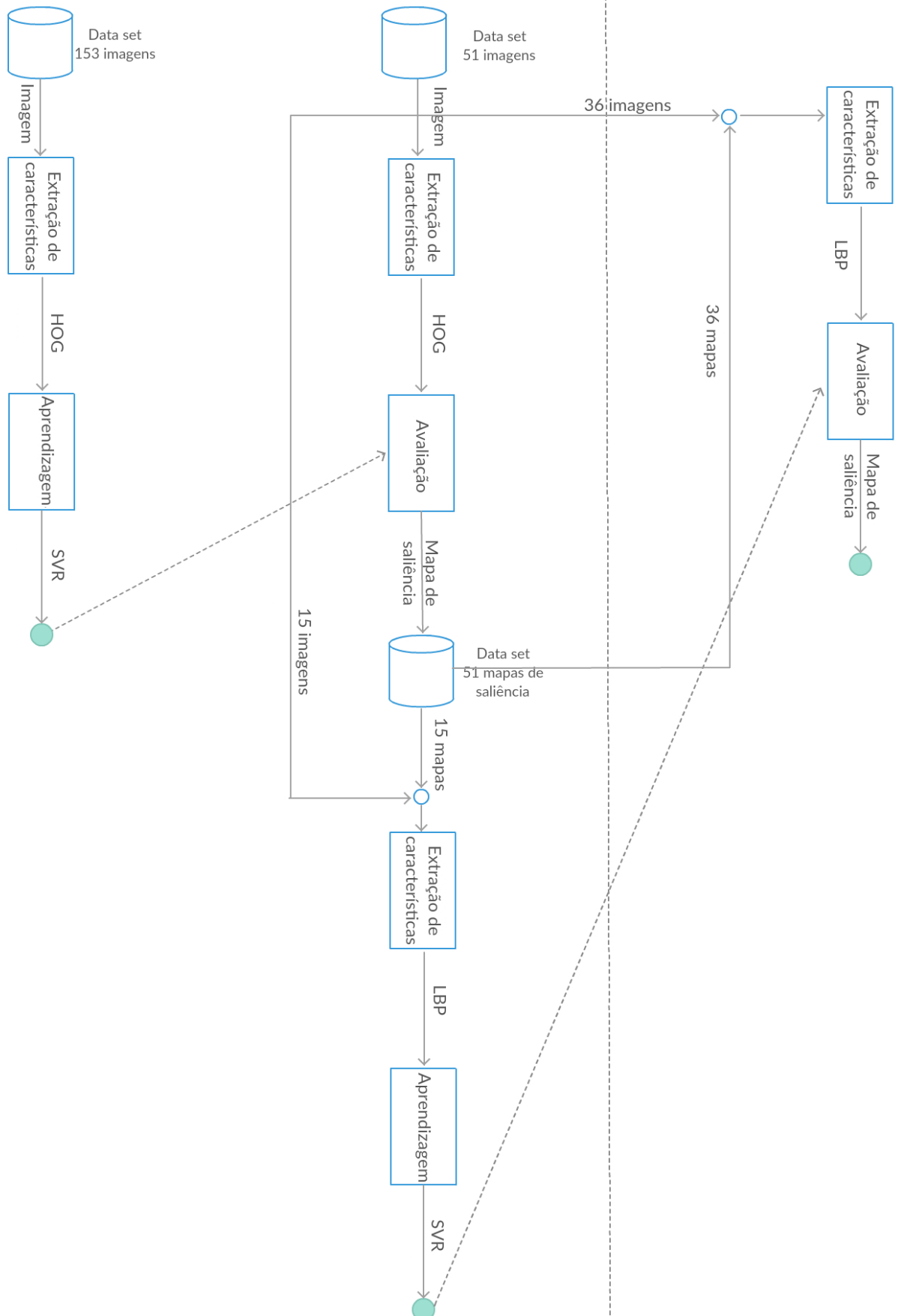
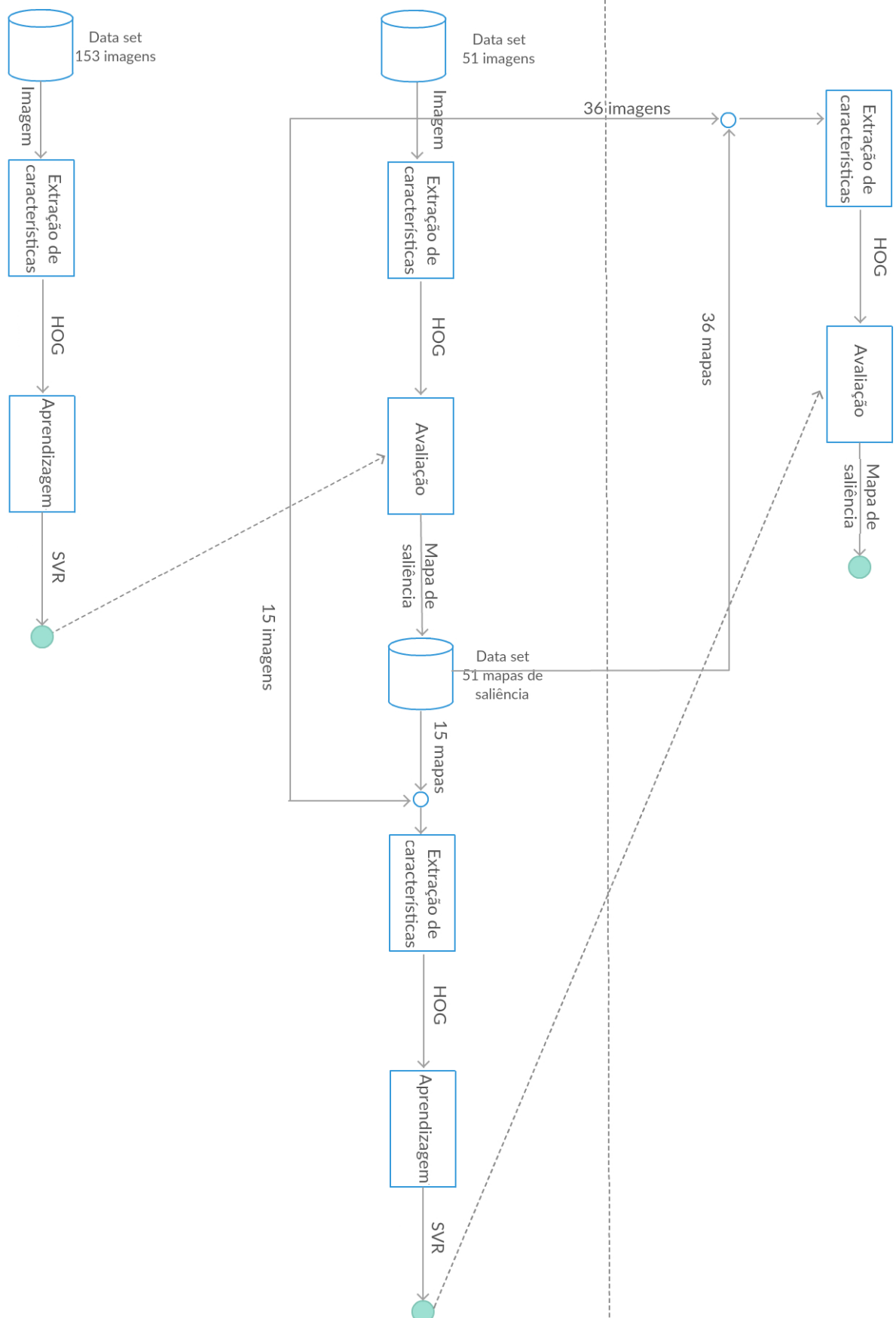


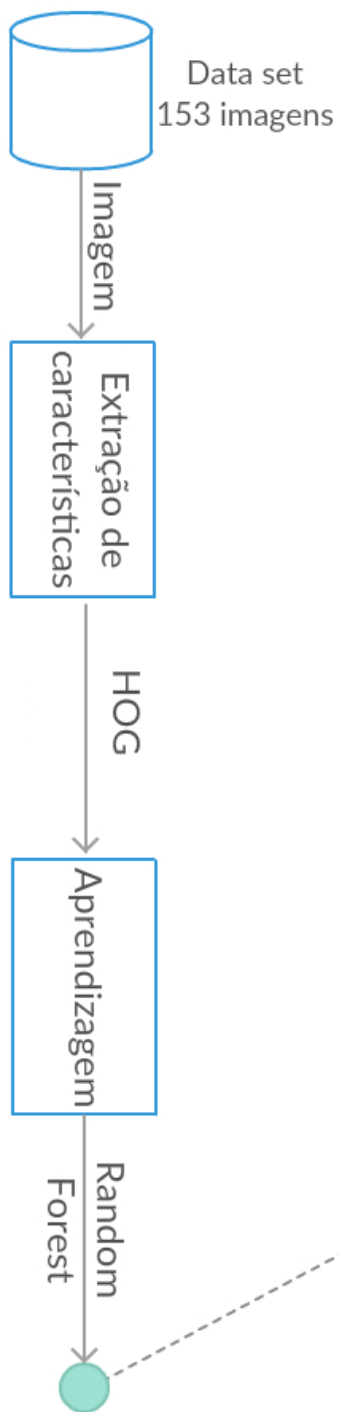
Figura 5.2: Diagrama do Experimento II

Etapa de treinamento

Etapa de teste



Etapa de treinamento



Etapa de teste

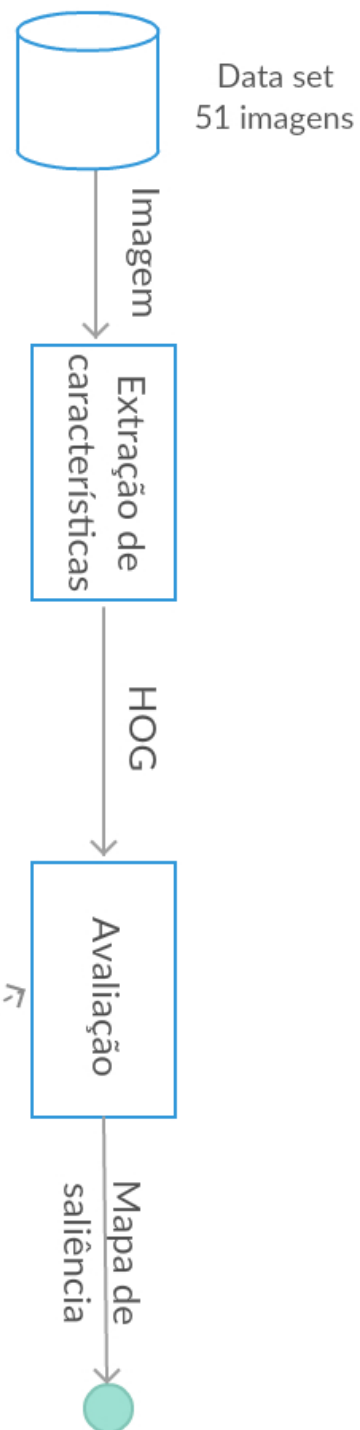


Figura 5.4: Diagrama do Experimento IV

Nos experimentos foram realizados testes utilizando 51 imagens da base de imagens *MIT Street Scene* com veículos em posição frontal já segmentados, como apresentado na Figura 5.5.



Figura 5.5: Amostras da base MIT Street Scenes

5.1 Avaliação dos Classificadores SVM

Foram construídos seis classificadores SVM sendo um para cada ponto, cada classificador foi treinado com 6885 amostras, sendo 1377 positivas (que contém o ponto de interesse centralizado na amostra) e 5508 negativas (obtidas aleatoriamente pelas imagens de treinamento onde o ponto de controle não está presente). O desempenho das máquinas de vetores de suporte foi auferido com base nas medidas de acurácia, sensibilidade, precisão e medida-F, métricas comumente empregadas para avaliar modelos preditivos [Faceli et al., 2011]. Estas medidas poderão exprimir o quão bom ou ruim foi o desempenho dos classificadores SVM, que tipo de erro (erro do tipo I — falso positivo ou erro do tipo II — falso negativo) foi mais frequente e o quanto o desempenho do reconhecedor de forma pode ser afetado pelo emprego desta abordagem de obtenção de pontos de controle.

Para obter estas medidas as saídas dos classificadores SVM, que originalmente foram treinadas como modelos de regressão, precisaram ser apresentadas como valores binários a fim de que as máquinas pudessem ser avaliadas como modelos preditivos. Tendo em vista que o mapa de saliência produzido pela SVM origina poucos valores distantes de 0 ou 1, conforme pode ser visto no histograma da Figura 5.6, utilizou-se o valor 0,55 como limiar para discretização das saídas. Para cada ponto de controle foram utilizadas 36 imagens de teste onde os rótulos foram marcados manualmente em cada uma delas.

Na Tabela 5.1 são apresentadas as medidas de desempenho para a classificação das SVM. Pode-se observar que a acurácia foi superior a 99%, indicando que em uma imagem a grande maioria dos *pixels* são classificados corretamente. No entanto, tendo em vista que apenas uma pequena seção contém a região de interesse, o erro cometido precisa ser avaliado também com base em outras medidas, particularmente no que diz respeito à classificação desta região. Comparando os modelos com base na precisão observa-se que a classificação dos pontos 1, 5 e 6 foi superior aos resultados dos pontos 2, 3 e 4, ou seja, nos modelos dos pontos 2, 3 e 4 a presença de exemplos pertencentes a classe negativa classificados como positivos foi mais frequente, o que impacta negativamente no alinhamento dos pontos de controle.

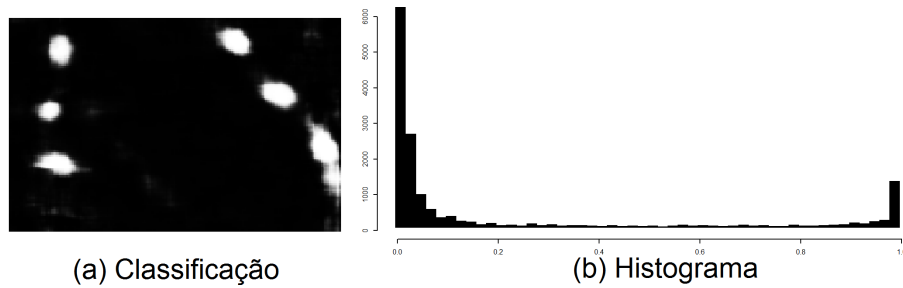


Figura 5.6: Histograma do mapa produzido pela SVM de regressão

Na prática, ainda que o ASM permita estabelecer dinamicamente o alinhamento da forma, permitindo lidar com problemas de oclusão, rotação e escala, a baixa precisão é a principal razão pela qual, em alguns casos, a forma do veículo deixa de ser detetada. Portanto, o problema mais desafiador é encontrar uma medida-F elevada, que represente um equilíbrio entre a sensibilidade e a precisão.

Tabela 5.1: Experimento I - Medidas de desempenho das SVM no primeiro nível

Medida	Ponto 1	Ponto 2	Ponto 3	Ponto 4	Ponto 5	Ponto 6
Acurácia	99,7%	99,7%	99,7%	99,7%	99,7%	99,7%
Precisão	63,1%	57,2%	56,9%	51,6%	61,3%	63%
Sensibilidade	71%	63,6%	75,4%	52,6%	68,3%	71,7%
Medida-F	0,668	0,602	0,649	0,521	0,646	0,671

Para reduzir a taxa de falso positivo foi utilizado um segundo nível de classificação. A ideia de usar uma cascata de classificadores tem como principal objetivo otimizar o reconhecimento dos veículos. A cascata de classificadores divide-se em dois estágios, a principal função do estágio inicial é classificar as regiões da imagem como verdadeiras ou falsas, no segundo estágio são consideradas apenas as regiões classificadas como positivas, assim o estágio mais elevado torna-se mais preciso para descartar falsos positivos na região que está sendo analisada.

No segundo estágio, 15 mapas de saliência produzidos pelo primeiro classificador foram utilizados para treinar 6 diferentes máquinas SVM, as mesmas imagens de teste utilizadas na medição do Experimento I também foram as mesmas dos demais experimentos. Dois experimentos foram realizados com diferentes características, o primeiro com LBP, Tabela 5.2, e o segundo com características HOG, Tabela 5.3. Quanto à utilização da cascata de classificadores, os resultados obtidos possuem maior precisão em ambos os casos em comparação com o experimento do primeiro estágio, a utilização do descritor HOG apresentou taxa de precisão superior ao uso do LBP na maioria dos casos, no entanto a combinação dos dois tipos de características foi mais eficiente na

localização de formas, em particular, forma de veículos como indica a Tabela 5.4.

Tabela 5.2: Experimento II - Medidas de desempenho das SVM no segundo nível com LBP

Medida	Ponto 1	Ponto 2	Ponto 3	Ponto 4	Ponto 5	Ponto 6
Acurácia	99,3%	99,3%	99,4%	99,6%	99,2%	99,4%
Precisão	71,1%	78,5%	79,9%	69,8%	82,9%	89,3%
Sensibilidade	7,9%	7%	33,3%	8%	7,8%	29,4%
Medida-F	0,142	0,129	0,47	0,144	0,142	0,443

Tabela 5.3: Experimento III - Medidas de desempenho das SVM no segundo nível com HOG

Medida	Ponto 1	Ponto 2	Ponto 3	Ponto 4	Ponto 5	Ponto 6
Acurácia	99,3%	99,3%	99,3%	99,6%	99,2%	99,2%
Precisão	85,8%	85,1%	88%	69,8%	97,1%	82,3%
Sensibilidade	6,1%	4,1%	7,9%	4%	4,1%	6,7%
Medida-F	0,111	0,078	0,144	0,076	0,078	0,124

Por último, para comparar os resultados obtidos, foram realizados testes com o classificador Florestas Aleatórias, do inglês *Random Forest* — RF. Assim como nos experimentos anteriores, cada ponto possui um classificador específico. Os resultados podem ser vistos na Tabela 5.4. Pode-se observar que os resultados de ambas as técnicas no primeiro estágio foram semelhantes, no entanto, a cascata de classificadores ainda se mostra mais precisa. Para analisar de forma mais geral, é necessário obter os resultados da utilização do ASM em todos os experimentos.

Tabela 5.4: Experimento IV - Medidas de desempenho das RF no primeiro nível

Medida	Ponto 1	Ponto 2	Ponto 3	Ponto 4	Ponto 5	Ponto 6
Acurácia	99,4%	99,3%	99,4%	99,6%	99,3%	99,4%
Precisão	75,5%	52,5%	73,3%	57,6%	66,3%	67%
Sensibilidade	35,5%	49,8%	39,7%	28,3%	23,6%	47%
Medida-F	0,483	0,511	0,515	0,38	0,348	0,552

5.2 Avaliação do Modelo de Forma Ativa

Para mensurar quantitativamente o desempenho dos modelos de formas ativas, foi utilizada a métrica empregada por Zhao et al. [2004], em que se calcula o erro através da distância euclidiana pela equação:

$$E = \frac{1}{N} \sum_{i=1}^N \left(\frac{1}{n} \sum_{j=1}^n \sqrt{(x_{ij} - x'_{ij})^2 + (y_{ij} - y'_{ij})^2} \right) \quad (5.1)$$

onde N é o número total de imagens de teste e n é o número de pontos de controle em uma imagem. (x_{ij}, y_{ij}) corresponde a localização do j -ésimo ponto de controle rotulado manualmente da i -ésima imagem de teste e (x'_{ij}, y'_{ij}) corresponde a localização do j -ésimo ponto de controle calculado. Assim como no trabalho de Zhao et al. [2004], neste também foi calculado o erro através da distância euclidiana e a taxa de melhoria dos demais experimentos em relação ao primeiro é calculada por:

$$M = \frac{E_{\text{experimentoI}} - E_{\text{experimentoII}}}{E_{\text{experimentoI}}} \times 100 \quad (5.2)$$

Tabela 5.5: Medidas de desempenho do ASM

Abordagem	Erro Médio	Melhoria
Experimento I	8,770%	—
Experimento II	6,434%	26,636%
Experimento III	6,669%	23,956%
Experimento IV	8,04%	8,323%

De acordo com a Tabela 5.5, fica evidente que a cascata de classificadores se destaca em relação ao primeiro nível de classificação, seja usando Máquinas de Vetores de Suporte ou Florestas Aleatórias. O refinamento da máquina permite que os *pixels* sejam melhores avaliados, alguns *pixels* classificados como positivos no primeiro nível foram classificados como negativos pelo segundo, devido a isso, a taxa de falsos positivos decaiu. Alguns resultados obtidos pelo Modelo de Forma Ativa em cada experimento são exibidos na Figura 5.7.

Uma das contribuições deste trabalho é a detecção de veículos parcialmente oclusos, na maioria dos casos em que houve oclusão, a abordagem conseguiu localizar a forma, Figura 5.8. Como a forma é rígida e se desloca de modo restrito, se a maior parte dos pontos são detetados corretamente então a forma irá se deslocar até encontrar o objeto desejado.

No trabalho de Li, Gu e Kanade [2011] que serviu de inspiração para a realização deste, foram usadas 1162 imagens com vista frontal da base *MIT Street Scene* para o desenvolvimento da abordagem, dessas imagens, 400 foram usadas no treinamento e o resto para teste. A raiz do erro médio quadrático (RMSE) de cada abordagem é mostrada nas Figuras 5.9 e 5.10, os pontos utilizados por Li, Gu e Kanade [2011] não foram os mesmo deste trabalho, considerando que o Experimento II foi mais bem sucedido entre os experimentos realizados neste trabalho e comparando com o trabalho base, este apresentou erros menores em 4 dos seis pontos em comum, além disso, a maioria dos veículo parcialmente oclusos foram detetados utilizando apenas o ASM enquanto no trabalho de Li, Gu e Kanade [2011] foi criado um novo modelo, denominado modelo de alucinação, que emprega Inferência Bayesiana para detetar as partes oclusas.

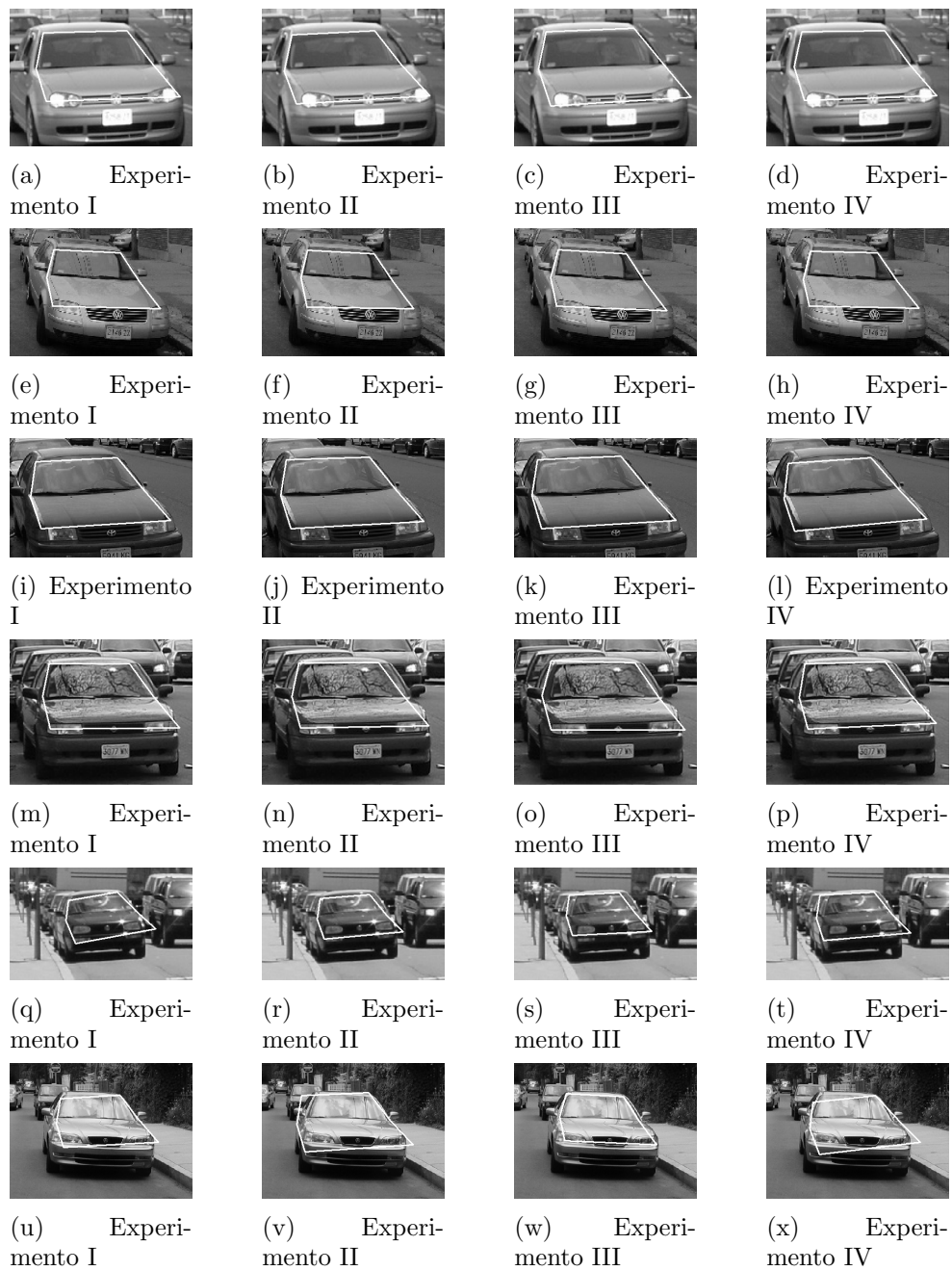


Figura 5.7: Resultados ASM em todos os experimentos

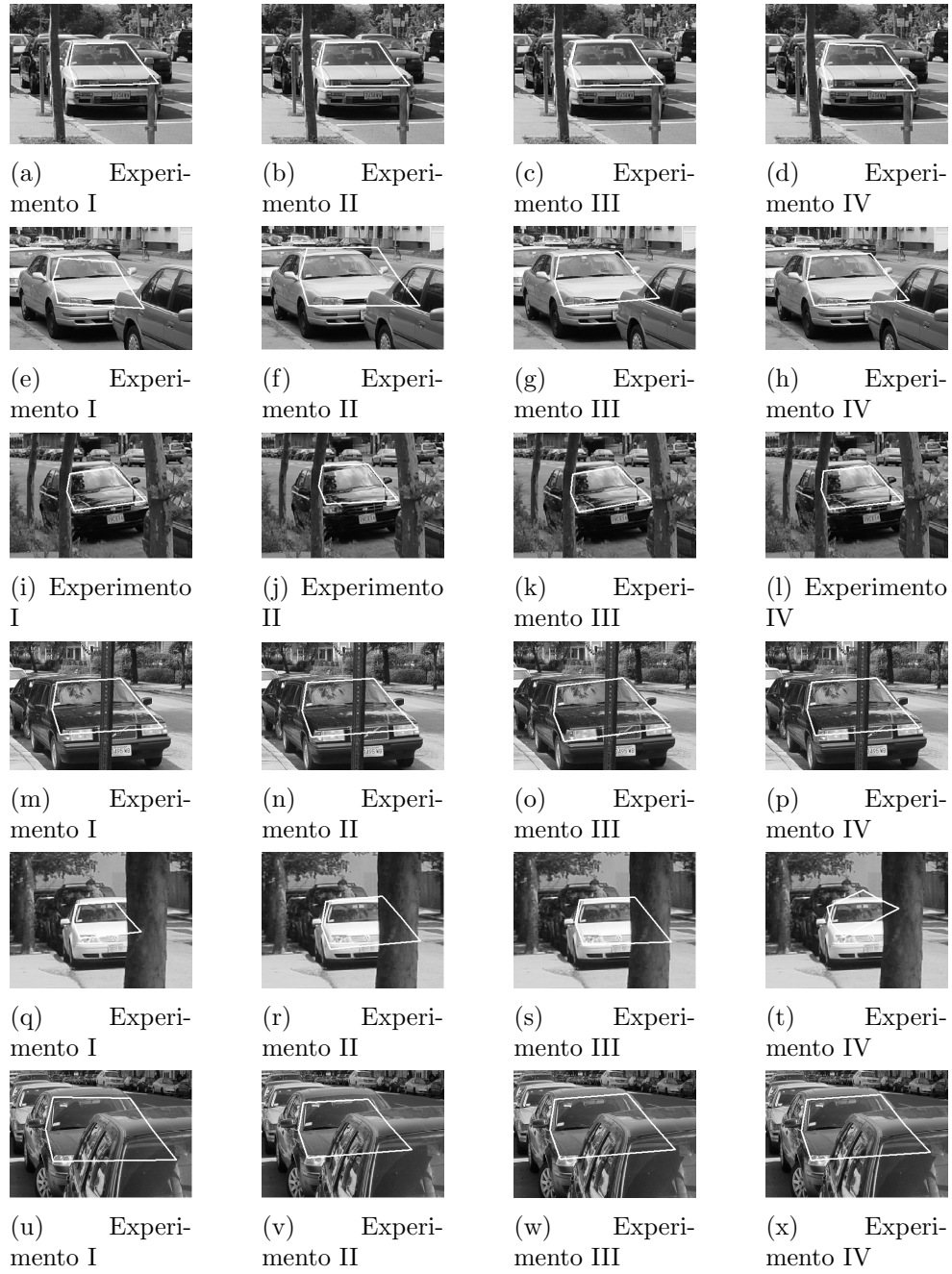


Figura 5.8: Resultados ASM em veículos oclusos

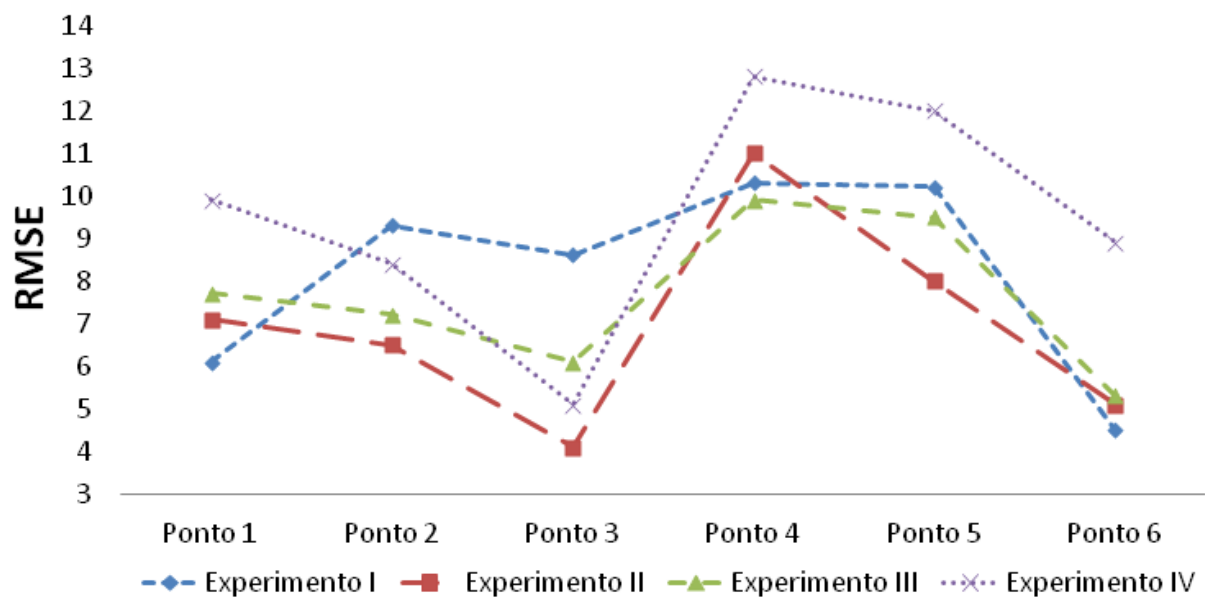


Figura 5.9: Erros dos experimentos I ao IV

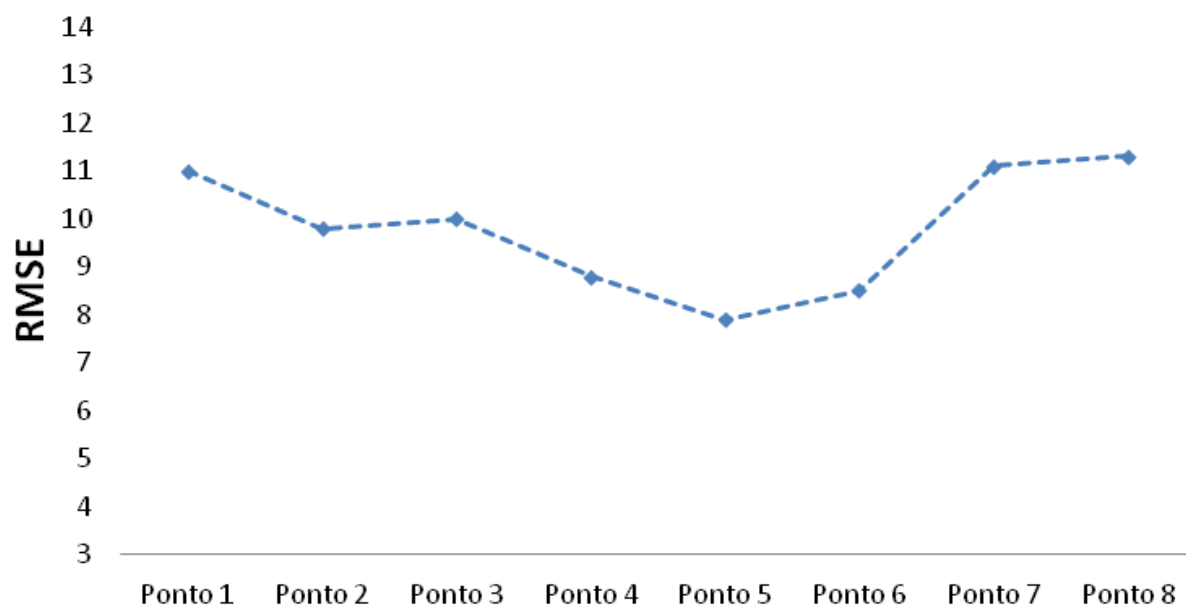


Figura 5.10: Erros da abordagem de Li, Gu e Kanade [2011] com vista frontal

Capítulo 6

Conclusões

Localizar veículos em imagens é uma tarefa bastante desafiadora para os algoritmos computacionais. A localização torna-se ainda mais difícil quando as imagens contêm veículos oclusos, no entanto, devido aos avanços na visão computacional podemos combinar métodos tornando-os mais complexos capazes de detetar veículos em diferentes situações. Neste trabalho foram estudados alguns aspectos teóricos referentes à detecção de veículos. Realizou-se uma revisão do estado da arte, sobretudo nas técnicas de aprendizagem de máquina e modelos deformáveis que viabilizou a implementação deste sistema.

Foi desenvolvida uma abordagem híbrida para identificar formas de veículos em imagens de alta dimensionalidade utilizando classificadores SVM para localizar pontos de controle em *patches* de 40×40 *pixels* e a classificação da forma com ASM, abordagem comumente utilizada para classificar formas de órgãos, expressões faciais ou outras formas ativas. Esta foi uma maneira de balancear a complexidade para solucionar um problema difícil distribuindo-a entre diferentes máquinas de aprendizado — a segmentação dos pontos de controle e o reconhecimento da sua forma geométrica. Também fez parte desse trabalho utilizar uma cascata de máquinas de vetores de suporte de forma a tentar melhorar o desempenho dos classificadores dos pontos de controle. Os *pixels* classificados pela primeira máquina como positivos foram submetidos ao segundo nível de classificação a fim de diminuir a quantidade de falsos positivos.

Os experimentos foram realizados com a base *MIT Street Scene* que contém mais de 3000 imagens de veículos. Dentre os métodos avaliados, as características de gradiente foram mais discriminantes para a tarefa de reconhecimento de veículos, além disso, os melhores resultados da classificação dos vetores de características foram obtidos com o uso de uma cascata de classificadores. Com relação aos modelos deformáveis, a combinação de aprendizagem de máquinas com modelos deformáveis resultou na principal contribuição deste trabalho: detecção de veículos parcialmente oclusos. Ainda que os pontos não tenham sido detetados pelas máquinas SVM, o modelo ASM permitiu o ajuste da forma ao veículo ocluso.

Neste trabalho, o erro da localização da forma de veículos foi de 6,4%, além disso, o uso dos modelos deformáveis solucionou grande parte das oclusões permitindo que a forma se alinhasse corretamente ao contorno. Isso confirma a hipótese de que a detecção do veículo, obtida pela composição de uma abordagem híbrida, com diferentes máquinas, proporcione melhores resultados.

É evidente que o uso das *Random Forest* não foi tão eficiente quanto na abordagem desenvolvida por Li, Gu e Kanade [2011] devido ao processo de treinamento não ter sido o mesmo, seja pela menor quantidade de imagens ou pela diferente configuração nos parâmetros da máquina. A abordagem desenvolvida neste trabalho detetou diversos pontos oclusos utilizando apenas modelos deformáveis, no entanto, no trabalho que serviu de inspiração para a realização deste, o modelo pontual foi enriquecido com um modelo de alucinação para resolver o problema da oclusão. Outro ponto importante do trabalho de Li, Gu e Kanade [2011] foi a detecção e alinhamento da forma em imagens de veículos com diferentes poses, isso tornou o sistema capaz de localizar veículos em diferentes situações, enquanto neste são detetados veículos apenas em posição frontal.

6.1 Trabalhos Futuros

Um trabalho científico nunca se encerra em si mesmo, pois sempre há chance de estendê-lo e melhorá-lo. Há algumas linhas de investigação que parecem promissoras, sugestivas de serem adotadas na continuação deste trabalho. Tendo em vista que a abordagem desenvolvida foi voltada para imagens estáticas, uma extensão natural que pode-se levantar é aplicação das técnicas aqui desenvolvidas para vídeo e realização de tomadas de decisão em tempo real. Mais recentemente, com a popularização de processamento paralelo em placas gráficas, diversos trabalhos, a exemplo de Kumar et al. [2016], Mahmoudi et al. [2016] e Yavari et al. [2015], apresentaram implementações de sistemas de processamento de vídeo em GPU, pois o processamento de matrizes, estrutura de dados que representam imagens, é naturalmente paralelizável.

A detecção de veículos consiste em localizar uma janela que representa a área em torno do veículo, um método muito utilizado para segmentar objetos é o de Viola e Jones [2001]. A partir da segmentação da janela contendo o veículo é possível realizar o alinhamento da forma, assim a integração da abordagem desenvolvida ao detetor de veículos permitirá que veículos sejam localizados automaticamente. Outro trabalho a ser desenvolvido é o alinhamento da forma em diferentes poses do veículo como feito por Li, Gu e Kanade [2011] possibilitando a realização de um estudo comparativo entre os dois trabalhos.

Referências Bibliográficas

- Acharya A., Lee J. e Chen A., *Real Time Car Detection and Tracking in Mobile Devices*, In Connected Vehicles and Expo (ICCVE), International Conference on. IEEE, pp. 239-240, 2012.
- Baumberg A. e Hogg D., *Learning Flexible Models from Image Sequences*, 3rd ed. European Conference on Computer Vision, Sweden, Stockholm, 1994.
- Brett A. D., Hill A., e Taylor C. J., *A Method of 3D Surface Correspondence for Automated Landmark Generation*, 8th British Machine Vision Conference, Essex, England, 1997.
- Cootes T. F., Taylor C. J., Cooper D. H. e Graham J., *Training Models of Shape from Sets of Examples*, Department of Medical Biophysics, University of Manchester, 1992.
- Cootes T. F. e Taylor C. J., *Active Shape Model Search using Local Grey-Level Models: A Quantitative Evaluation*, British Machine Vision Conference, BMVA Press: 639/648, 1993.
- Cootes T. F., Taylor C. J., e Lanitis A., *Active Shape Models: Evaluation of a Multi-Resolution Method for Improving Image Search*, British Machine Vision Conference, BMVA, 1994.
- Cootes T. F., Taylor C. J., Cooper D. H. e Graham J., *Active Shape Models – Their Training and Application*, Computer Vision and Image Understanding, v. 61, n. 1, pp. 38-59, 1995.
- Cootes T. F. e Taylor C. J., *Statistical Models of Appearance for Computer Vision*, Imaging Science and Biomedical Engineering, University of Manchester, 2004.
- Dalal N. e Triggs B., *Histograms of Oriented Gradients for Human Detection*, Proc. IEEE CS Conf. Computer Vision and Pattern Recognition, vol. 1, pp. 886-893, 2005.
- Faceli K., Lorena A. C., Gama J. e Carvalho A. P. L. F., *Inteligência Artificial: uma Abordagem de Aprendizado de Máquina*, LTC, Rio de Janeiro, 2011.

- Fernández-Delgado M., Cernadas E., Barro S. e Amorim D., *Do we need hundreds of classifiers to solve real world classification problems?*, The Journal of Machine Learning Research, 15(1), pp. 3133-3181, 2014.
- Freund Y. e Schapire R. E., *A Decision-Theoretic Generalization of On-Line Learning and an Application to Boosting*, In EuroCOLT, Springer-Verlag, pp. 2337, 1995.
- Gil A. C., *Como elaborar projetos de pesquisa*, 5. ed. São Paulo: Atlas, 2008.
- Gower J. C., *Generalized Procrustes Analysis*, Psychometrika, v. 40, n. 1, pp. 33-51, 1975.
- Hadi R. A., Sulong G. and George L. E., *Vehicle Detection and Tracking Techniques: A Concise Review*, Signal and Image Processing: An International Journal, Vol.5, 2014.
- Hill A., e Taylor C. J., *Automatic Landmark Generation for Point Distribution Models*, Fifth British Machine Vision Conference, England, York, BMVA Press, 1994.
- Kass M., Witkin A. e Terzopoulos D., *Snakes: Active Contour Models*, International journal of computer vision, v. 1, n. 4, pp. 321-331, 1988.
- Kumar P. , Singhal A., Mehta S. e Mittal A., *Real-time moving object detection algorithm on high-resolution videos using GPUs*, Journal of Real-Time Image Processing, pp. 93-109, 2016.
- Kuo C. H. e Nevatia R., *Robust Multi-View Car Detection Using Unsupervised Sub-Categorization*, Applications of Computer Vision, Workshop on. IEEE, 2009.
- Li Y., Gu L. e Kanade T., *Robustly Aligning a Shape Model and its Application to Car Alignment of Unknown Pose*, Pattern Analysis and Machine Intelligence, IEEE Transactions, v. 33, n. 9, pp. 1860-1876, 2011.
- Lichman M., *UCI Machine Learning Repository* [<http://archive.ics.uci.edu/ml/>], Irvine, CA: University of California, School of Information and Computer Science, 2013.
- Lin H. T., Lin C. J. e Weng R. C., *A Note on Platt's Probabilistic Outputs for Support Vector Machines*, Machine learning, v. 68, n. 3, pp. 267-276, 2007.
- Liu N., e Lovell B. C., *Hand Gesture Extraction by Active Shape Model*, Proc. Digital Image Computing: Techniques and Applications, pp. 59-64, 2005.
- Ma J. e Ren F., *Detect and Track the Dynamic Deformation Human Body With the Active Shape Model Modified by Motion Vectors*, IEEE International Conference on Beijing, pp. 587- 591, 2011.

- Madhogaria S., Baggenstoss P. M., Schikora M., Koch W. e Cremers D., *Car Detection by Fusion of HOG and Causal MRF*, Aerospace and Electronic Systems, IEEE Transactions, v. 51, n. 1, pp. 575-590, 2015.
- Mahmoudi S. A., Ammar M., Joris G. L. e Abbou A., *Real Time GPU-Based Segmentation and Tracking of the Left Ventricle on 2D Echocardiography*, In Bioinformatics and Biomedical Engineering, Springer International Publishing, pp. 602-614, 2016.
- Platt J., *Probabilistic Outputs for Support Vector Machines and Comparisons to Regularized Likelihood Methods*, Advances in large margin classifiers, v. 10, n. 3, pp. 61-74, 1999.
- Ratajczak R., Grajek T., Wegner K., Klimaszewski K., Kurc M. e Domanski M., *Vehicle Dimensions Estimation Scheme Using AAM on Stereoscopic Video*, In Advanced Video and Signal Based Surveillance (AVSS), 2013 10th IEEE International Conference on. IEEE, pp. 478-482, 2013.
- Tsai M. K., Lin Y. L., Hsu W. H. e Chen C. W., *Content-based vehicle retrieval using 3d model and part information*, In Acoustics, Speech and Signal Processing (ICASSP), IEEE International Conference, pp. 1025-1028, 2012.
- Tsin Y., Genc Y. and Ramesh V., *Explicit 3D modeling for vehicle monitoring in non-overlapping cameras*, In Advanced Video and Signal Based Surveillance, AVSS'09, Sixth IEEE International Conference, pp. 110-115, 2009.
- Vasconcelos M. J., Tavaré S. e Ribeiro J. M., *Modelos Pontuais de Distribuição em Visão Computacional*, In: 06 Congresso Nacional Mecânica Experimental, 2005.
- Viola P. e Jones M., *Rapid Object Detection Using a Boosted Cascade of Simple Features*, Conference on Computer Vision and Pattern Recognition (CVPR), Vol. 1, IEEE Computer Society, pp. 511-518, 2001.
- Wang S. J., Mathew A., Chen Y., Xi L. F., Ma L. e Lee J., *Empirical analysis of support vector machine ensemble classifiers*, Expert Systems with Applications, pp. 6466-6476, 2009.
- Wimmer A., Soza G. e Hornegger J., *A Generic Probabilistic Active Shape Model for Organ Segmentation*, in Proc. Int. Conf. Med. Image Comput. Comput. Assist. Intervent., pp. 26-33, 2009.
- Yavari M. , Li M., Li S. e Qi M., *Parallelizing abnormal event detection in crowded scenes with GPU*, In Fuzzy Systems and Knowledge Discovery (FSKD), 12th International Conference, pp. 274-277, 2015.

- Zang W., Wu Q. M. J., Yang X. e Fang X., *Multilevel Framework to Detect and Handle Vehicle Occlusion*, Intelligent Transportation Systems, IEEE Transactions, vol. 9, pp. 161-174, 2008.
- Zhao S., Gao W., Shan S. e Yin B., *Enhance the alignment accuracy of active shape models using elastic graph matching*. In Biometric Authentication, Springer Berlin Heidelberg, pp. 52-58, 2004.
- Zia M. Z., Stark M. e Schindler K., *Explicit Occlusion Modeling for 3d Object Class Representations*, Computer Vision and Pattern Recognition, IEEE Conference, pp. 3326-3333, 2013.